IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 8 May 2023, accepted 17 May 2023, date of publication 23 May 2023, date of current version 30 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3279022

RESEARCH ARTICLE

Generative Adversarial Network-Based Signal Inpainting for Automatic Modulation Classification

SUBIN LEE[1], YOUNG-IL YOON[2], AND YONG JU JUNG[1], (Member, IEEE)
[1]School of Computing, Gachon University, Seongnam 13120, Republic of Korea
[2]Research and Development Center, LIG Nex1, Seongnam 13488, Republic of Korea

Corresponding author: Yong Ju Jung (yjung@gachon.ac.kr)

This work was supported by the Korea Research Institute for Defense Technology Planning and Advancement (KRIT) Grant funded by the Defence Acquisition Program Administration (DAPA) under Grant KRIT-CT-22-002.

ABSTRACT Automatic modulation classification (AMC) aims to automatically identify the modulation type of a detected signal in an intelligent wireless receiver, such as software-defined radio (SDR). Recently, deep learning-based methods such as convolutional neural networks have been applied to AMC, showing high-accuracy performance. However, the earlier studies do not consider various signal degradations that can possibly occur during the transmission and reception of wireless signals. Particularly, the signal reception can be often unstable, and the signal can be partially received due to the dynamic spectrum sensing or signal sensing in the intelligent wireless systems. The corrupted signal with missing samples considerably degrades the accuracy of modulation classification of the deep learning-based models, because it is very different from the training datasets. To address this issue, the preprocessing process of restoring the corrupted signal, called signal inpainting, is essential. Although it is significant for the modulation classification, no studies have been performed to investigate the effect of signal inpainting on AMC. To that end, this study proposes a generative adversarial network(GAN)-based signal inpainting method that fills in the missing samples in a wireless signal. The proposed inpainting method can restore the time-domain signal with up to 50% missing samples while maintaining the global structure of each modulation type. The correct recovery of the global structure enables the extraction of distinctive features that play a key role in the modulation classification. To investigate this effect of signal inpainting on AMC, we perform intensive experiments on the RadioML dataset that has been widely used in the AMC studies. We compare the accuracy performance of the two state-of-the-art AMC models with and without the proposed signal inpainting, respectively. Through the analysis of the results, we show that the proposed GAN-based inpainting method significantly improves the accuracy of AMC.

INDEX TERMS Signal inpainting, signal restoration, modulation classification, CNN, GAN, SDR.

I. INTRODUCTION

The usage of wireless communications is widespread in all facets of daily life, the military, and industry. However, the problem of frequency resource depletion is arising as numerous devices such as the Internet of Things (IoT) and smartphones compete for wireless communications. This problem has led to active research on cognitive radio [1]. A cognitive radio can be configured dynamically to use the best vacant channels, which intelligently detects available channels in wireless spectrum to avoid user interference. As such, modern communication systems with intelligent spectrum management are being applied in military and civilian communication fields. In particular, automatic modulation classification (AMC) is a core technology in many software-defined radio (SDR) platforms that identifies the modulation type of nearby radio signals to prevent radio interference and improve spectrum efficiency [2].

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

In early studies in the field of wireless communications, AMC was studied based on traditional pattern recognition methods. Traditional approaches can be broadly divided into two categories: likelihood-based [3] and feature-based methods [4]. Likelihood-based methods use Bayesian inference for modulation classification by assuming prior information such as channel and noise models. However, likelihood-based estimate methods suffer from high computational complexity and they are not appropriate for dynamically changing channel conditions. Feature-based methods extract handcrafted features based on predefined rules and use them to classify modulation types using a machine learning classifier such as support vector machine. However, the extraction of good features in a handcrafted manner, which achieves high classification accuracy, is challenging. These feature-based approaches have the limitation of low classification accuracy due to their inability to handle the complexity of modern communication systems. On the other hand, the most recent approaches use deep learning (DL)-based models that can perform AMC without requiring any prior knowledge.

DL-based AMC methods performs feature extraction and classification simultaneously in an end-to-end model, enabling high accuracy in complex classification problems that are difficult to address with traditional handcrafted methods. Recent studies have shown that DL models such as convolutional neural networks (CNNs) can effectively extract features from the various representation of wireless signals such as in-phase and quadrature (IQ) signal or spectrogram to achieve high accuracy in the modulation classification [5], [6], [7], [8]. However, DL-based methods generally require a large amount of training data, and the classification accuracy depends on the quality of the training data [9]. To collect the training data for AMC, various characteristics of communication channels must be considered in complex real environments, which can result in high costs due to the need for various modulation types and labeling. Recently, simulation data considering various communication environments have been generated through SDR systems and used for training the AMC model. Although the use of simulation data has partially resolved the lack of data issue, it has still limitations in that it does not consider all the characteristics that can occur in real-world communication environments. By this reason, if it is applied to real communication systems, the classification accuracy of AMC can be considerably reduced, as mentioned in previous studies [10], [11].

In real wireless communication environments, various deteriorations can occur such as attenuation due to distance, noise, multipath fading, and partial loss of transmitted data. The RadioML dataset [12], [13], [14], which has been commonly used in recent AMC studies, has been designed to simulate various channel characteristics such as noise, attenuation, and fading similar to real wireless communications. However, it does not consider the data loss that can possibly occur during wireless communications. As mentioned in previous studies [10], [11], the signal reception can be often unstable and the signal can be partially received due to the dynamic spectrum sensing or signal sensing in the SDR platforms. That is, burst signals, received during short durations with uncertain starting and ending time in the dynamic spectrum access, can significantly degrade the accuracy of AMC. Hence, an advanced method is needed to recognize the modulation type accurately even when only a part of the wireless signal data is received in various conditions. Particularly, a preprocessing step is required to restore partially received data before the modulation classification (see Figure 1).

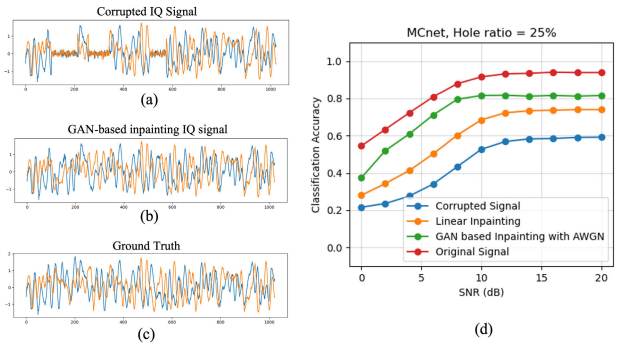


FIGURE 1. IQ signal inpainting for automatic modulation classification in an intelligent wireless receiver. (a) Corrupted IQ signal with missing samples and AWGN in 128-QAM modulation. (b) Restored signal obtained by the proposed GAN-based inpainting method. (c) Ground-truth signal. (d) Accuracy of the modulation classification by MCnet [7]. Note that the classification accuracy is significantly improved by applying the proposed inpainting method before the modulation classification.

Signal inpainting aims at restoring degraded or missing samples in a one-dimensional or two-dimensional signal. In the computer vision field, many studies have been recently conducted on the image inpainting task using DL-based methods [15], [16], [17], [18], [19], [20]. Particularly, many CNN-based inpainting methods have been proposed in the last decade, and the state-of-the-art methods for image inpainting are mainly based on the generative adversarial networks (GANs) [21]. GANs are generative models that can generate realistic images in image synthesis tasks such as image inpainting and super-resolution. Notably, in previous studies of wireless communications, GAN has been also used for data augmentation to effectively train an AMC model, and it was reported that the synthetic wireless signal generated through GAN is very similar to real data [22], [23], [24].

Despite the compelling performance shown by recent DL models for two-dimensional image inpainting, few research has been dedicated to developing the inpainting methods for one-dimensional signals. Among them, notable research [25], [26] is the audio signal inpainting that aims to restore lost audio signals during transmission. However, no studies have been conducted for wireless signal inpainting in wireless communications. Moreover, it is necessary to study the effect of signal inpainting on AMC in the situations where only a part of the wireless signal is received.

This study proposes a DL-based signal inpainting method that restores the signal loss in wireless communications. More importantly, we also conduct intensive experiments to investigate the effect of signal inpainting on AMC. Specifically,

we propose a GAN-based signal inpainting model that consists of a Unet-style generator network and a discriminator network. The generator fills in the missing samples of a corrupted IQ signal. The restored signal is then used for the modulation classification using a backbone AMC model. The contributions of this study are as follows:

- A GAN-based signal inpainting method is proposed to restore the corrupted signal with missing samples in wireless communications. The proposed inpainting method can restore the time-domain signal while maintaining the global structure of each modulation type. The correct recovery of the global structure enables the extraction of distinctive features that play key role in the modulation classification.
- To further investigate the effect of the signal inpainting on the automatic modulation classification, we perform intensive experiments with the RadioML dataset. Specifically, the restored signal is used for two state-of-the-art models of CNN-based AMC (i.e., RobustCNN [8] and MCnet [7]). Then, we compare the modulation classification accuracy between the models with and without signal inpainting. The experimental results indicate that the signal inpainting can improve the classification accuracy by an average of 11% to 34% for various signals in a SNR range of 0dB to 20dB.

The remaining structure of this paper is organized as follows. Section II presents related works. Section III describes the proposed GAN-based signal inpainting method. Section IV describes the experimental results on the effect of signal inpainting on AMC. A conclusion is presented in Section V.

II. RELATED WORK

A. DEEP LEARNING-BASED MODULATION CLASSIFICATION

Deep learning is now being used in many industries to solve difficult recognition tasks. DL is also widely utilized in the wireless communication fields. In the AMC task, various DL techniques such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have been recently applied to improve the classification accuracy performance.

RNN/LSTM can effectively capture long-term dependencies of a sequence data in time domain [27], [28]. Since a radio signal is represented by I and Q signals in the time series, RNN/LSTM has achieved high classification accuracy performance in AMC. Rajendran et al. [29] proposed a DL-based classification model using a 2-layer LSTM network for the feature extraction and a fully connected layer for the classification. They showed that their model outperforms traditional approaches using SVMs and random forest classifiers. Another study [30] attempted to detect wideband spectra using an LSTM-based classifier with a temporal attention layer to handle various channel impairments. Chen et al. [31] proposed a single-layer LSTM network with a random erasing-based test time augmentation mechanism (RETTA) and an attention module to improve the classification accuracy. They evaluated the effectiveness of their method on the RadioML2016.10A dataset [13] and outperformed simple RNN and LSTM-based baseline approaches. Meanwhile, Huang et al. [32] also attempted to solve the overfitting problem by adopting various data augmentation methods including rotation, flipping, and Gaussian noise addition for both IQ and AP samples in the training and prediction stages of a LSTM-based classifier.

Many studies [5], [6] have also used CNN models for the modulation classification task because CNN can effectively extract deep features from continuous signals. O'Shea et al. [14] proposed two CNN models based on VGG [33] and ResNet [34] architectures using 1D asymmetric convolutional filters. Also, MCnet [7] is a CNN model designed for robustly identifying the modulation type in various channel impairment environments. It effectively reduces the number of parameters by using vertical and horizontal 1D asymmetric filters in convolution blocks and enables efficient training the model by using various skip connections to solve gradient vanishing problems in convolution layers. For the RadioML 2018.01A dataset [14], it achieved high modulation classification accuracy with low complexity. A similar study, RobustCNN [8], also applied MCnet's asymmetric filter and skip connection techniques. The RobustCNN uses a frame replication method that flips a radio signal to extract features more effectively in convolution operations. The experimental results showed that it can achieve high classification accuracy. Huynh-The et al. [35] proposed Chain-Net that captures the meaningful information of modulation signal at multi-scale feature representations with various asymmetric kernels.

In AMC, although most studies have used IQ signals as input for their DL models, other techniques have also used different modalities such as constellation images. In addition, there have been attempts to represent a radio signal to a spectrogram image using the short time Fourier transform (STFT). Since this modulation classification is an image-based classification task, a study [36] investigated the AMC based on the off-the-shelf CNNs such as AlexNet [37] and GoogleNet [38]. Although these image-based methods can achieve high classification accuracy, they require high computational complexity and running time. To mitigate these drawbacks, lightweight CNN models using various constellation representations have also been proposed [39], [40]. In recent studies [41], [42], [43], transformer-based models have also been proposed to facilitate the long-term dependency of time-domain signals. For more details of the AMC methods, we can refer to a recent survey paper [11]. In this study, we use two state-of-the-art CNN models (i.e., RobustCNN [8] and MCnet [7]) as backbone AMC models whose performance has been verified on the RadioML 2018.01a dataset [14].

B. DEEP LEARNING-BASED SIGNAL INPAINTING

Signal inpainting aims to restore the degraded or missing samples of a signal. In case of two-dimensional signal

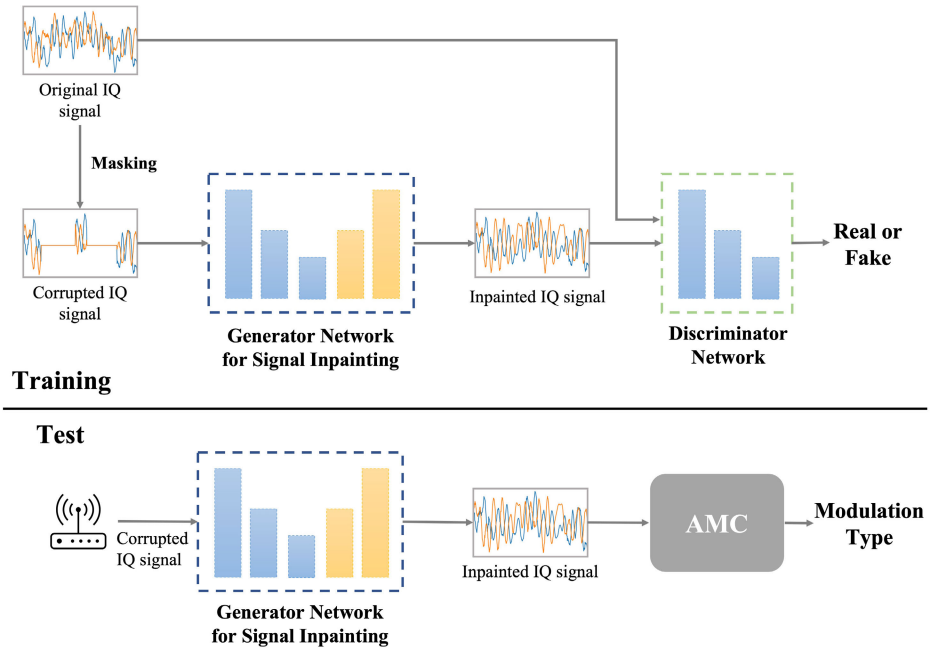


FIGURE 2. Overall architecture of the proposed GAN-based IQ signal inpainting network that consists of a generator and a discriminator for the adversarial training. Note that, in test time, the output signal of the generator is used for the modulation classification in the backbone AMC.

inpainting, image inpainting is one of the active research fields in computer vision, which fills in the missing pixels in an image. Many studies [15], [16], [17], [18], [19], [20] have proposed CNN-based image inpainting methods and the state-of-the-art methods are mostly based on GAN models. In a GAN, two neural networks (i.e., generator and discriminator) contest with each other, enabling the generation of realistic images [15], [16], [17], [18], [19], [20]. Although DL-based image inpainting has been well studied, one-dimensional signal inpainting has been rarely studied.

In case of one-dimensional signal inpainting, some previous studies [25], [26] have proposed audio inpainting methods that restores partially missing samples in an audio signal. Early audio inpainting studies [44], [45] attempted by extracting handcrafted features using statistical methods. These traditional methods showed good results for an audio signal in a short range (e.g., less than 50 ms) and repetitive patterns such as music, but they were not suitable for non-periodic signals such as speech. Accordingly, DL-based approaches have gained attention. Ebner and Eltelt [46] proposed a DL-based model using the Wasserstein GAN (WGAN). In their study, they investigated the restoration of audio signals with longer ranges of loss (500ms - 550ms) than the short losses (less than 50ms) studied in previous research. They established a generalized model for various types of instruments using WGAN, focusing on waveforms rather than spectrograms. Chang et al. [47] proposed a hybrid method using both spectrogram and waveform data for long-range losses. Also, a study [48] attempted to restore audio more accurately by adding video frame information.

As mentioned earlier, the wireless signal is often partially received and hence the signal inpainting enables better performance of signal recognition tasks by using the restored signals in an intelligent wireless receiver. Nevertheless, signal inpainting for restoring wireless communication signals has not been well researched due to their irregular patterns. More importantly, no studies have been performed to investigate the impact of signal inpainting on AMC.

It should be further mentioned that the adversarial attack is one of big issues in AMC. Recent studies have reported that deep learning AMC models are vulnerable to adversarial attacks [49]. To tackle this issue, recent studies have attempted to develop robust AMC models such as ensemble-based deep learning model [50] that can mitigate various adversarial attacks in AMC. Although such recent approach showed that robust AMC models can achieve significant degree of mitigation on malicious attacks, it is inevitable to reduce the accuracy of AMC to a great extent. As another approach for mitigating adversarial attacks in AMC, the signal restoration framework can be also used to restore maliciously corrupted signals and the restored signal is then used for a AMC model to withstand adversarial attacks such adversarial noise corruption. If we can construct a good training dataset that consists of pairs of original and attacked signals, the signal restoration approach will be beneficial to mitigate adversarial attacks in AMC. Furthermore, another possible approach is a hybrid method that uses both robust AMC model and signal restorations approaches in an ensemble manner. We just note that mitigating the adversarial attack is

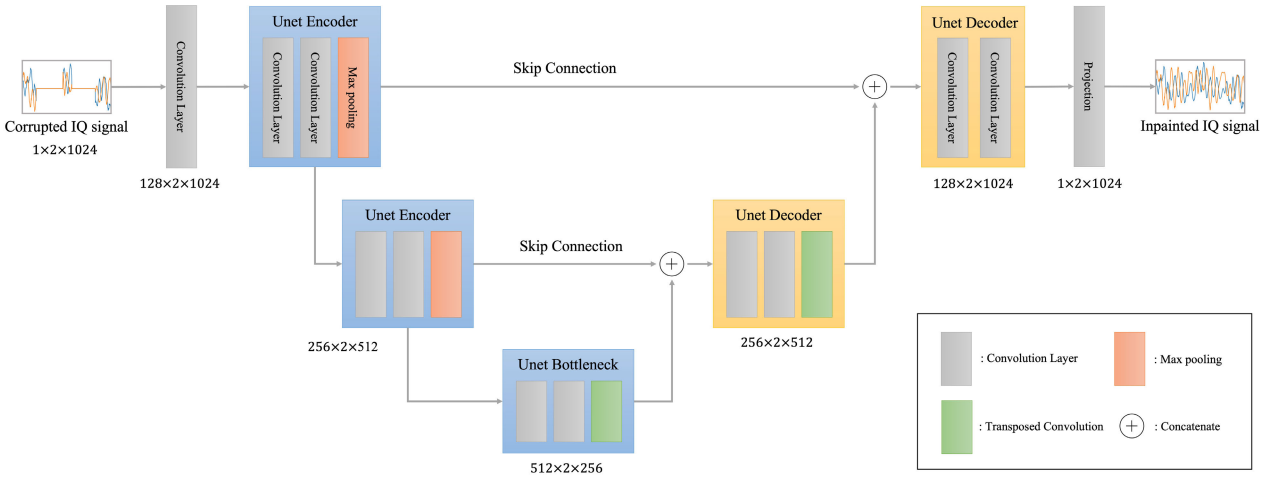


FIGURE 3. Structure of the generator network for IQ signal inpainting.

TABLE 1. Structure of the generator network.

Layer	Feature Shape	Description
Input	$1 \times 2 \times 1024$	IQ signal
Convolution	$128 \times 2 \times 1024$	kernel (3, 7), padding (1, 3)
Unet Encoder-1	$128 \times 2 \times 1024$	$2 \times$ [kenel (3, 7), padding (1, 3)]
Pooling	$128 \times 2 \times 512$	max pooling (1, 2)
Unet Encoder-2	$256 \times 2 \times 512$	$2 \times$ [kenel (3, 7), padding (1, 3)]
Pooling	$256 \times 2 \times 256$	max pooling (1, 2)
Unet Bottleneck	$256 \times 2 \times 256$	$2 \times$ [kenel (3, 7), padding (1, 3)]
Transposed Convolution	$256 \times 2 \times 512$	kernel (1, 2), stride (1, 2)
Concatenate	$512 \times 2 \times 512$	concat(Transposed Convolution, Unet Encoder-2)
Unet Decoder-2	$256 \times 2 \times 512$	$2 \times$ [kenel (3, 7), padding (1, 3)]
Transposed Convolution	$128 \times 2 \times 1024$	kernel (1, 2), stride (1, 2)
Concatenate	$256 \times 2 \times 1024$	concat(Transposed Convolution, Unet Encoder-1)
Unet Decoder-1	$128 \times 2 \times 1024$	$2 \times$ [kenel (3, 7), padding (1, 3)]
Projection	$1 \times 2 \times 1024$	kernel (1, 1)

out of scope of our study and remains as an important future research.

III. GAN-BASED IQ SIGNAL INPAINTING

GAN [21] consists of two basic components: a generator and a discriminator. In the original GAN, the generator is trained to map from the latent space to a distribution that is close to the real data distribution, while the discriminator is trained to distinguish between the real data distribution and the generated fake data distribution. The generator and the discriminator are trained through the adversarial learning, which means that the generator is trained to fool the discriminator. Through this adversarial learning technique, the generator can generate the synthetic data that is similar to the real data. Figure 2 shows that architecture of the proposed IQ signal inpainting network based on GAN.

For the training of the proposed GAN model, we simulate the corrupted signal with random loss in which there are arbitrary missing samples at random locations in a signal. Specifically, we use the masking technique that randomly removes some samples in an original IQ signal to make a masked signal with hole. Hereafter, we call the hole as missing samples in a signal. Note that, in previous studies [15], [16], [17], [18], [19], [20], this masking technique has been widely used to train the image inpainting models. A detailed description of the masking technique used in this study is described in Section IV-A1 The masked signal is then used as input for the generator network, as shown in Fig. 2. The generator fills in the hole in the input time-domain signal through several convolution layers. The restored IQ signal is then used as an input to the discriminator along with the original signal. The discriminator determines whether the input signal is a real signal or a fake signal. Through the adversarial learning, the generator is trained to generate a fake signal with the same realistic characteristics as the original IQ signals.

During the testing, the generator is used to perform signal inpainting on a corrupted input signal with missing samples at arbitrary locations. The restored IQ signal is then passed into a backbone AMC model that classifies it into one of the modulation types. In the following sections, we describe details of the generator and discriminator networks.

A. GENERATOR NETWORK FOR SIGNAL INPAINTING

The generator network has a Unet-like structure [51] that consists of the encoder and decoder layers in a hierarchical

TABLE 2. Structure of the discriminator network.

Layer	Feature Shape	Description
Input	$2 \times 2 \times 1024$	IQ signal and inpainted IQ signal
Convolution	$128 \times 2 \times 1024$	kernel (3, 7), padding (1, 3)
Pooling	$128 \times 2 \times 512$	max pooling (1, 2)
Convolution	$256 \times 2 \times 512$	kernel (3, 7), padding (1, 3)
Pooling	$256 \times 1 \times 256$	max pooling (2, 2)
Convolution	$256 \times 1 \times 256$	kernel (1, 5), padding (0, 2)
Pooling	$256 \times 1 \times 128$	max pooling (1, 2)
Convolution	$256 \times 1 \times 128$	kernel (1, 5), padding (0, 2)
Global Average Pooling	$32 \times 1 \times 128$	
Dense Layer	256	dense layer (8196×256)
Dense Layer	1	dense layer (256, 1)

manner. Figure 3 shows the architecture of the generator network. Note that the Unet-like architecture is mainly used for the image-to-image synthesis tasks in computer vision. The network's input is a corrupted IQ signal. In this study, we use the RadioML 2018.01A dataset [14] for training and testing the model. In this dataset, each signal has 1024 samples. Therefore, an input IQ signal has a size of $1 \times 2 \times 1024$ ($channel \times height \times width$).

One of the most important design factors in a CNN-based signal inpainting is to increase the receptive field. Due to the locality characteristics of the convolution operation, each feature is extracted in a local area that a convolutional kernel works. In general, constructing a deeper network can increase the receptive field [15], [16], [17], [52]. However, deeper network can cause the overfitting problem. Therefore, the proposed network maintains a shallow network with three spatial levels while increasing the kernel size as 3×7, allowing us to increase the receptive field. Figure 3 shows the proposed generator network for IQ signal inpainting.

All convolutions in the generator use a 3×7 kernel with a padding size of 1×3. Batch normalization [53] and ReLU activation functions follow each convolution layer. In the first convolution layer, the input IQ signal $\mathbf{X}_{input} \in \mathbb{R}^{(1\times2\times1024)}$ is transformed into a feature $\mathbf{F} \in \mathbb{R}^{128\times2\times1024}$.

In the encoder, each level consists of two convolution layers, followed by a max pooling with a 1×2 kernel to reduce the width in the horizontal direction. The feature ${\mathbf{F}_{Encoder}}^{(i)}$ in the i^{th} layer is given by

$$\begin{aligned}&{\mathbf{F}_{Encoder}}^{(i)}\\&= MaxPool(ConvLayer(ConvLayer({\mathbf{F}_{Encoder}}^{(i-1)}))),\end{aligned} \tag{1}$$

$$\begin{aligned}&ConvLayer(\mathbf{F})\\&= RELU(BN(conv(\mathbf{F}))),\end{aligned} \tag{2}$$

where BN denotes the batch normalization.

While reducing the width of the feature by half using max pooling, the channel size is doubled in the following convolution. For example, the feature with a size of $128 \times 2 \times 512$ obtained after passing through the first level of the encoder is increased to $256 \times 2 \times 512$ feature by the convolution in the next level. In the bottleneck, the width is also reduced by half and the channel is doubled through a convolution operation. After two convolution layers in the bottleneck, a transposed convolution is performed to increase the width of the feature.

In the decoder, the width of the feature is also doubled through a transposed convolution. The kernel size of the transposed convolution is 1×2 and the stride is 1×2. Also, a skip connection is used to maintain the low-level information extracted from the encoder. The transposed convolution's output and the corresponding encoder feature at the same level, passed through the skip connection, are concatenated and then passed into a convolution layer.

$$\begin{aligned}{\mathbf{F}_{skip}}^{(i)} = Concat(&{\mathbf{F}_{Encoder}}^{(i)},\\&TransConv({\mathbf{F}_{Decoder}}^{(i+1)})),\end{aligned} \tag{3}$$

$${\mathbf{F}_{Decoder}}^{(i)} = ConvLayer(ConvLayer({\mathbf{F}_{skip}}^{(i)})), \tag{4}$$

where $TransConv()$ denotes the transposed convolution.

By following this approach, the feature is restored to the original size of $128 \times 2 \times 1024$. Finally, the last projection layer projects the feature into the same dimension as the original input (i.e., $1 \times 2 \times 1024$). A detailed structure of the generator is shown in Table 1.

B. DISCRIMINATOR NETWORK

The discriminator determines whether the inpainted signal is fake or real. The network structure of the discriminator is shown in Fig. 4. The discriminator concatenates the restored signal and the original as an input (size of $2 \times 2 \times 1024$). Similar to a typical classifier, it passes through convolution layers and performs max pooling operations to ultimately perform a binary classification between real and fake. In the training process, the ReLU in the convolutional layers is replaced with LeakyReLU to effectively transfer the gradients generated by the discriminator to the generator [54].

Specifically, the discriminator generates features by passing through four convolution layers and performing a global average pooling [55]. Finally, it performs the classification of real and fake through two fully connected layers. The initial two convolution layers generate features using a 3×5 kernel, and the subsequent two convolutions use a 1×5 kernel. The detailed structure of the discriminator is presented in Table 2.

C. LOSS FUNCTION

In a traditional GAN [21], the adversarial loss function is defined by

$$\begin{aligned}&\min_G \max_D \mathcal{L}_{GAN}(G, D)\\&= \mathbb{E}_{\mathbf{y}\sim p_{data}(\mathbf{y})}[\log D(\mathbf{y})] + \mathbb{E}_{\mathbf{x}\sim p_x(\mathbf{x})}[\log(1 - D(G(\mathbf{x})))],\end{aligned} \tag{5}$$

where $\mathcal{L}_{GAN}(G, D)$ is the combined loss for the GAN structure. $D(\mathbf{y})$ is the discriminator's result of predicting the real or fake, $G(\mathbf{x})$ is the fake data generated by the generator, and $D(G(\mathbf{x}))$ is the discriminator's result of predicting the fake data.

From the discriminator's perspective, it is trained to distinguish between the synthesized data from the generator and

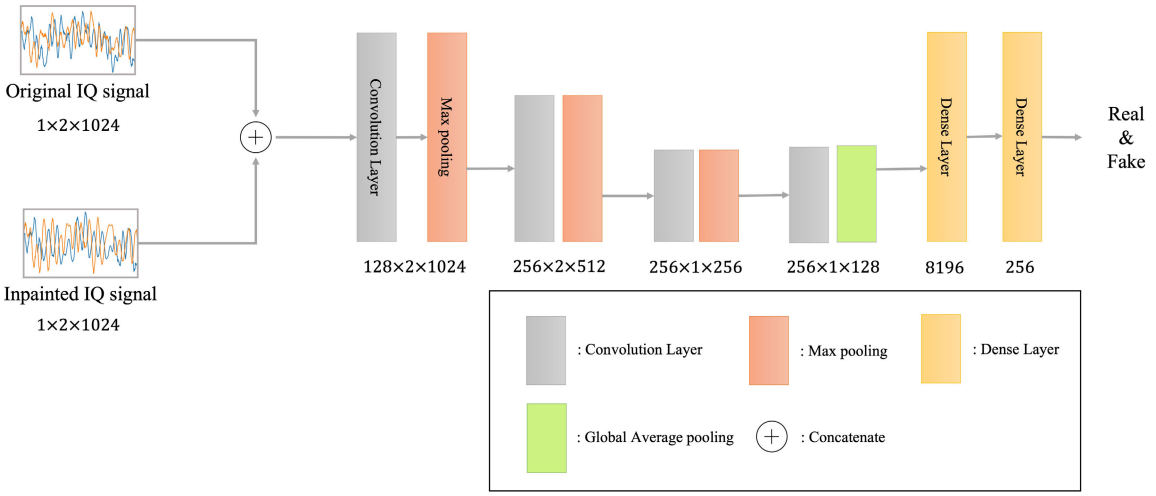


FIGURE 4. Structure of the discriminator.

the real data. That is, it is trained to maximize $\mathcal{L}_{GAN}(G, D)$. On the other hand, the generator is trained to generate data that the discriminator cannot distinguish between real and generated data, so it is trained to minimize $log(1 - D(G(\mathbf{x})))$.

In IQ signal reconstruction, L1 loss is used to restore the corrupted signal as close to the original signal as possible. The L1 loss is given by

$$\mathcal{L}_{L1}(G) = |\mathbf{y}_{GT} - (G(\mathbf{x}))|, \tag{6}$$

where y_{GT} is the ground-truth signal.

The total loss is obtained by the weighted sum of the adversarial loss and signal reconstruction loss.

$$G^* = \text{argmin}_G \text{max}_D[\mathcal{L}_{GAN}(G, D) + \lambda \mathcal{L}_{L1}(G)], \tag{7}$$

Note that $\lambda = 10$ was used in our implementation.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENT SETTING

1) DATASET

As used in previous AMC studies, we used the RadioML 2018.01A dataset [14] in this study. Note that the RadioML dataset, generated using GNU Radio [12], is widely used for training AMC models. This dataset was created by considering many radio channel characteristics close to real radio signals through a SDR system. Table 3 shows the specification of this dataset. The dataset includes 24 modulation types (BPSK, QPSK, 8-PSK, 16-PSK, 32-PSK, 4-ASK, 8-ASK, 16-APSK, 32-APSK, 64-APSK, 128-APSK, 16-QAM, 32-QAM, 64-QAM, 128-QAM, 256-QAM, OOK, GMSK, OQPS, FM, AM-SSB-WC, AM-SSB-SC, AM-DSB-WC, and AM-DSB-SC). Each modulation type consists of various signals with SNRs ranging from −20dB to +30dB in 2dB intervals. Each signal has a size of 2×1024 and is composed of I and Q signals with a sample length of 1024. The dataset includes various channel effects such as multi-path fading, carrier frequency offset, symbol rate offset, delay spread, and additive white Gaussian noise (AWGN) to provide the simulated signals in a wide variety of communication environments. Since the accuracy of AMC is very low in extremely low SNR signals, in this experiment, the performance of the proposed method was measured in a SNR range of 0 dB to 20 dB.

2) IMPLEMENTATION DETAILS

The proposed signal inpainting model was trained on Nvidia RTX 3090 GPU with 24GB memory, using the Pytorch framework [56]. Adam optimizer [57] was used for training,

TABLE 3. RadioML 2018.01A dataset [14] specification.

Modulations	Analog	FM, AM-SSB-SC, AM-SSB-WC, AM-DSB-SC, AM-DSB-WC
	Digital	BPSK, QPSK, 8-PSK, 16-PSK, 32-PSK, 16-APSK, 32-APSK, 64-APSK, 128-APSK, 4-ASK, 8-ASK, 16-QAM, 32-QAM, 64-QAM, 128-QAM, 256-QAM, OOK, GMSK, OQPSK
SNR range	-20dB : +30 dB (2dB step)	
Signal shape	[2, 1024] (time-series, I / Q)	
Total number of signals	2,555,904 (4096 frames per modulation-SNR combination)	
Channel specification	Selective multi path Rician fading Carrier frequency offset Symbol rate offset Non-impulsive delay spread Doppler shift AWGN	

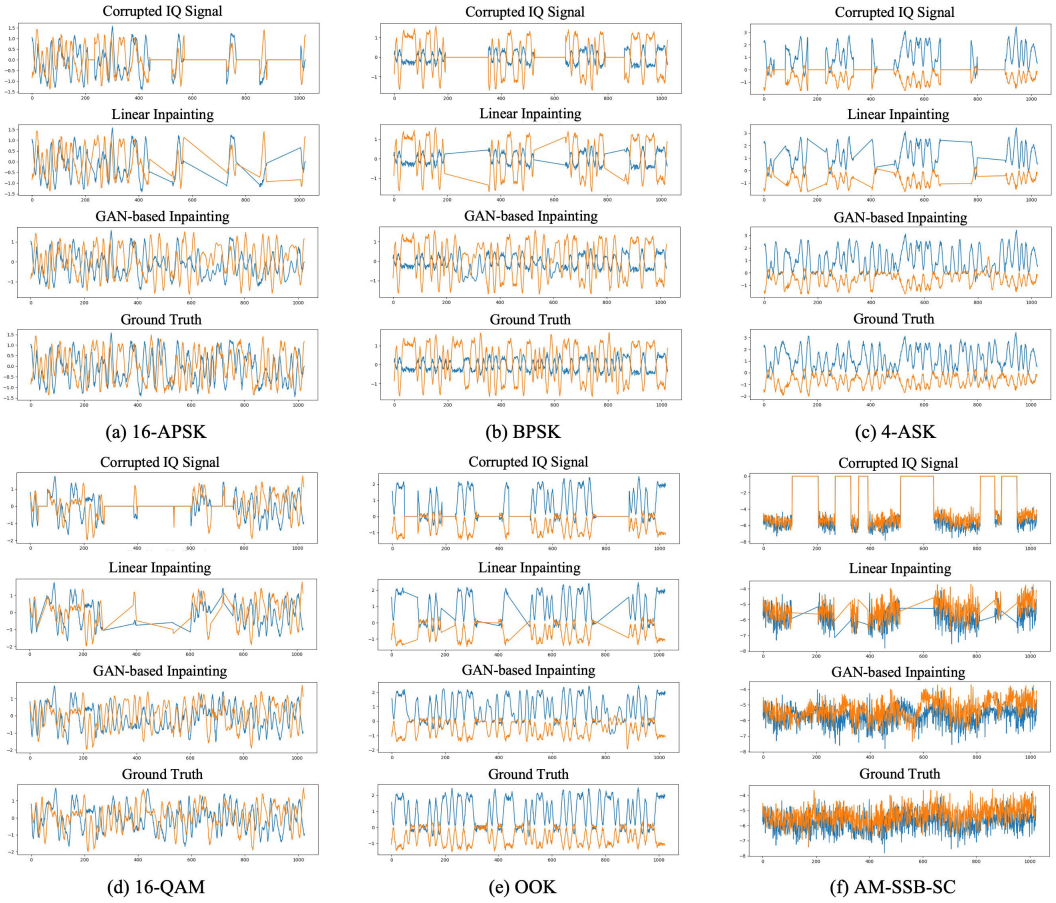


FIGURE 5. Visual results of signal inpainting. Note that, in each graph, blue color represents in-phase (I) component and orange color represents quadrature (Q) component of each time-domain signal, respectively.

with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $eps = 1e-08$ as optimizer parameters. The learning rate was set to 0.0001, and the training was conducted with a batch size of 128. A total of 100 epochs were trained.

3) IQ SIGNAL MASKING FOR THE HOLE GENERATION

The RadioML dataset used in this experiment was generated to closely mimic real wireless signals by considering many wireless channel effects, as shown in Table 3. However, it does not consider signal losses that possibly occur during the transmission process. In this experiment, a masking technique was used to simulate the corrupted signal with missing samples at arbitrary locations and sizes, as used in the image inpainting studies [15], [16], [17], [18], [19], [20].

Specifically, we first generated a random binary mask ($\mathbf{M}$) consisting of 0 and 1 for each sample location. Here 0 value represents a missing sample. This binary mask is element-wise multiplied with the original signal ($\mathbf{X}$) to create a masked signal ($\mathbf{X}_{masked}$) with some missing samples.

$$\mathbf{X}_{masked} = \mathbf{X} \otimes \mathbf{M}, \tag{8}$$

To simulate various lengths of missing samples, we randomly generate the holes (i.e., missing samples) with an arbitrary size ranging from 16 to 128 at random positions in an IQ signal. The length of the holes (i.e., the number of missing samples) in a signal is determined by the hole ratio parameter. In the training phase of our model, various hole ratios ranging from 12.5% to 50% were used. In the test, we used three different hole ratios (i.e., 12.5%, 25%, and 50%). For example, see Fig. 5 for the signals generated with 50% hole ratio. Here, 50% hole ratio means that 50% samples are the missing samples. Note that, in our experiments, the possible maximum size of a connected hole is given by $1024 \times holeratio$.

4) BACKBONE AMC MODELS

As described in Section III, the modulation classification is performed using the restored signals obtained by the proposed IQ signal inpainting method. Note that we can apply any AMC model that uses an IQ signal as an input. In our experiments, two state-of-the-art CNN models were used (i.e., MCnet [7] and RobustCNN [8]), which are publicly available. The classification results were analyzed to investigate the impact of signal inpainting on AMC. Specifically, we measured the average classification accuracy, individual accuracy per modulation type, and confusion matrices to analyze the effect of signal inpainting.

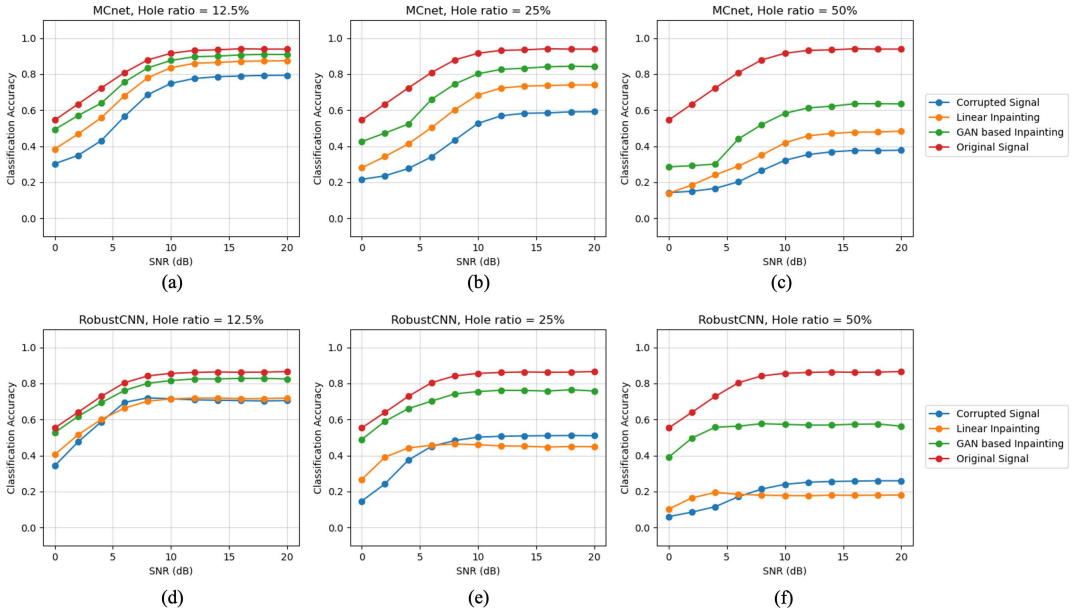


FIGURE 6. Accuracy of the modulation classification. Note that two state-of-the-art AMC models (i.e., MCnet [7] and RobustCNN [8]) have been used in our experiments. Each graph shows the accuracy result obtained by three different hole ratios (i.e., the percentage of missing samples in a signal).

B. EFFECT OF SIGNAL INPAINTING ON MODULATION CLASSIFICATION

1) VISUAL RESULTS OF SIGNAL INPAINTING

Figure 5 shows the restored IQ signals obtained from the experiments with a hole ratio of 50%, where the blue line represents the I component and the orange line represents the Q component of each time-domain signal. This figure shows examples of inpainting results for six different modulations. Note that the inpainting results for other modulation signals, which are not shown here due to space limitations, can be found in the supplementary material. In each figure, the graph of the first row shows the corrupted IQ signal with holes as input of the signal inpainting model, and the second row shows the result restored by simple linear interpolation. The third row shows the signal restored by the proposed GAN-based inpainting model, and the fourth row shows the ground truth signal.

The purpose of signal restoration is to generate a signal that is structurally similar to the real signal of a same modulation type. When comparing the restored signal with the ground-truth signal, it can be seen that the global structure of the generated signal is similar to that of the ground-truth signal. For example, when looking at the characteristics of 4-ASK and OOK signals in Figures. 5(c) and (e), I and Q components of the ground-truth signal have a symmetrical structure (i.e., I has positive values, while Q has negative values). We can observe the similar characteristics in the signals restored by the proposed method. Also, the AM-SSB-SC signal has an abruptly changing structure of I and Q components, as shown in Figure 5(f). The restored signal shows the same tendency of the global structure.

These results indicate that the proposed inpainting method can restore the IQ signals while maintaining the global structure of each modulation type. Intuitively, we can consider that the global structure of a time-domain signal plays key role as distinctive features in the modulation classification. Therefore, the correct recovery of the global structure will be very important for the automatic modulation classification, which enables the extraction of distinctive features for the modulation classification.

2) ACCURACY OF THE MODULATION CLASSIFICATION

Figure 6 shows the modulation classification accuracy measured by two state-of-the-art AMC models (i.e., RobustCNN [8] and MCnet [7]) after performing IQ signal inpainting with different hole ratios (i.e., 12.5%, 25%, and 50%). Each graph shows four different accuracy measurements: the original signals, corrupted signals with missing samples, signals restored by linear interpolation, signals restored by the proposed inpainting model. It can be observed that the AMC accuracy decreases when holes occur in the IQ signal compared to the original IQ signal, and the degree of accuracy decreases even more, as the hole ratio increases.

In the case of MCnet [7] (Figures. 6(a), (b) and (c)), the original signal achieved about 90% accuracy at an SNR of 10dB and up to 94% accuracy in higher SNR conditions. However, when the modulation classification was performed using the corrupted signals with holes, the accuracy decreased significantly. At hole ratio = 12.5% and SNR = 10dB (Figure 6(a)), the accuracy was about 80%, which is 10% lower

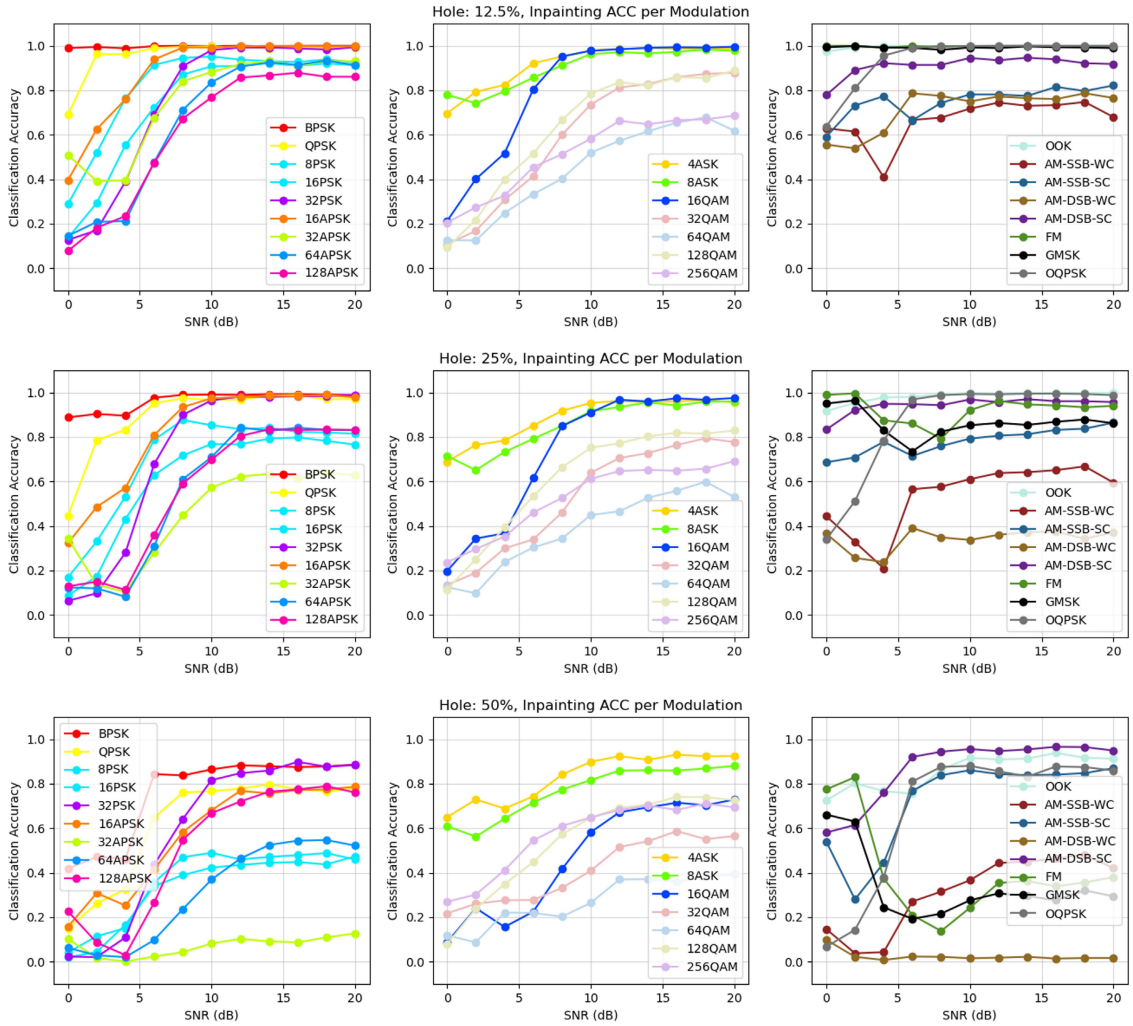


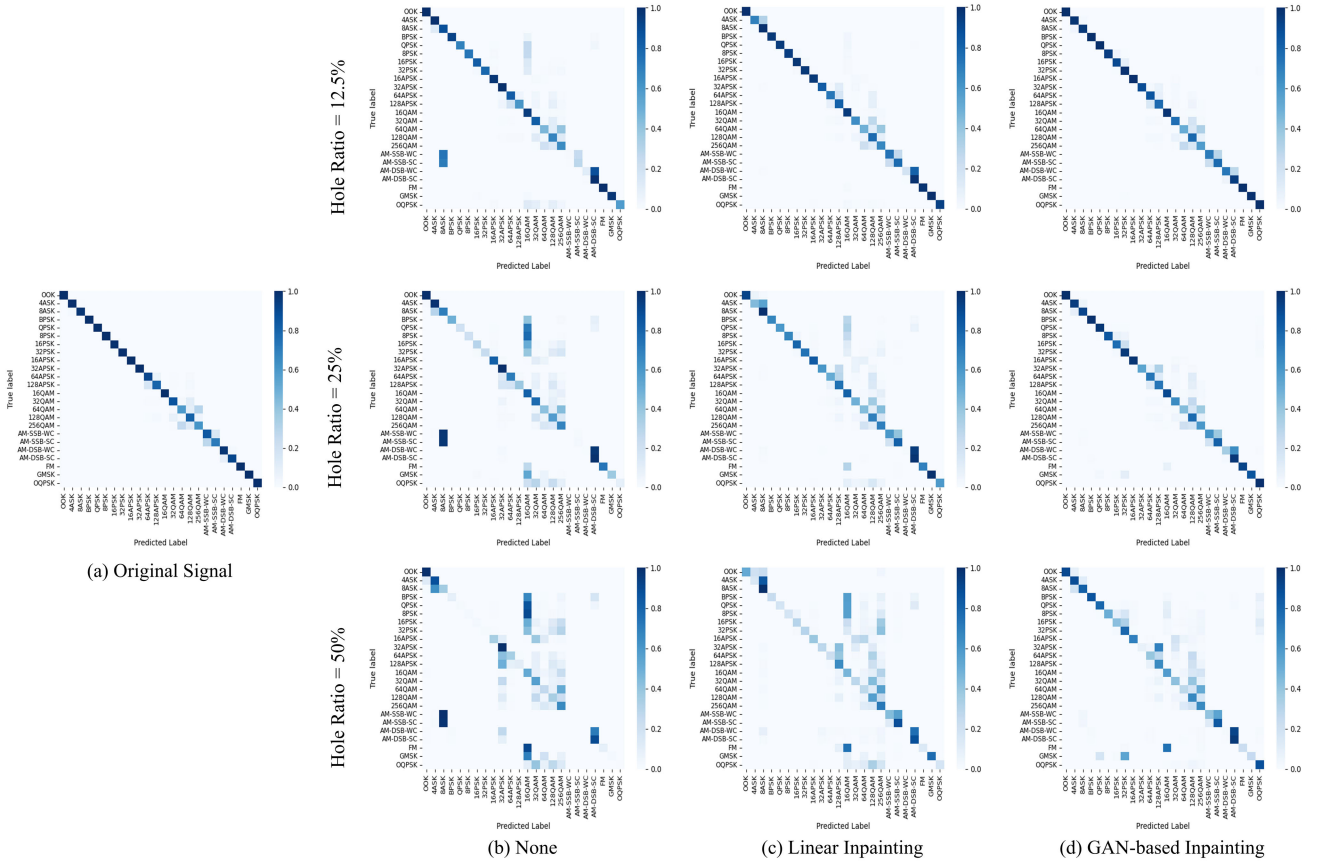
FIGURE 7. AMC accuracy per modulation type.

than that of the original signal. At hole ratio = 25% (Figure 6(b)), the accuracy decreased by 58%, and at hole ratio = 50% (Figure 6(c)), the accuracy dropped to about 37%. This is considered an inevitable performance decline as the deep learning-based AMC has been trained on the signals without holes (i.e., large difference between the training and testing datasets).

When holes were filled using the simple linear interpolation, the classification accuracy was better than not filling the holes, but the accuracy was still significantly reduced compared to the original signal. Particularly, when there were many holes, such as in the case of hole ratio = 50%, the performance drop was significant due to the distortion of the original signal. The accuracy was 87% for hole ratio = 12.5%, 74% for hole ratio = 25%, and 48% for hole ratio = 50%.

The proposed deep learning-based inpainting method shows that the classification accuracy is slightly lower by about 3% compared to the original signal. It achieves 91% accuracy at hole ratio = 12.5%. In addition, it shows a significantly higher accuracy of 84% at hole ratio = 25%, which is 25% higher than that of the corrupted signal. These results indicate that the proposed inpainting method can mitigate the performance degradation of AMC in situations where missing samples can occur.

Figures 6(d), (e), and (f) show the accuracy results of using another backbone AMC, the RobustCNN model [8]. In this AMC model, the accuracy of the corrupted signal decreases more significantly as the hole ratio increases compared to the original signal. While the original signal achieved an accuracy of 86%, the corrupted signal achieved 70% accuracy at hole ratio = 12.5%, 50% accuracy at hole ratio = 25%, and 25% accuracy at hole ratio = 50%, exhibiting a greater performance degradation. This indicates that the RobustCNN [8] is highly sensitive to the corrupted signals with holes. In addition, when the linear interpolation is applied to restore the corrupted signal, the classification accuracy is similar to the accuracy of the corrupted signal, showing 71% accuracy at



(a) Original Signal (b) None (c) Linear Inpainting (d) GAN-based Inpainting

FIGURE 8. Confusion matrix.

hole ratio = 12.5%, 47% accuracy at hole ratio = 25%, and 18% accuracy at hole ratio = 50%. This result indicates that applying linear interpolation lowers the performance. In contrast, the proposed inpainting model achieves 82% accuracy at hole ratio = 12.5%, 76% accuracy at hole ratio = 25%, and 57% accuracy at hole ratio = 50%. As a result, it can be seen that the proposed inpainting method can improve the accuracy performance of AMC.

3) ACCURACY ANALYSIS BY MODULATION TYPE

Figure 7 shows the accuracy graphs of the modulation classification for 24 modulation classes using the MCnet model. For the visualization purpose, the graphs show three categories into three groups: PSK+APSK, ASK+QAM, and low-order +analog modulations.

In the case of PSK and APSK modulations shown in the first column of Figure 7, the accuracy curves look like the curves of the average accuracy in Figure 4. At a hole ratio of 12.5%, the average accuracy is above 80%, and for relatively easy modulations such as BPSK, the accuracy is close to 100% regardless of SNRs. As the hole ratio increases, the accuracy decreases, and at a hole ratio of 50%, the classification of the high-order modulations becomes difficult.

Similarly, in the case of ASK and QAM modulations shown in the second column of Figure 7, as the hole increases the accuracy decreases. The classification of ASK modulations shows relatively higher accuracy regardless of the hole ratio. However, high-order QAM modulations show low accuracy, as the hole increases.

In the case of the low-order and analog modulations shown in the third column of Figure 7, relatively high accuracy is achieved for all modulations at a hole ratio of 12.5%. At a hole ratio of 25% even in low SNR conditions. All modulations except AM-SSB-WC and AM-DSB-WC achieve similar accuracy to that at a hole ratio of 12.5%. However, at a hole ratio of 50%, FM and GMSK modulations show a significant decrease in accuracy, and the AM-DSB-WC modulation is not properly classified. In the next chapter, a confusion matrix analysis will be performed to further analyze the reasons for these phenomena.

4) ANALYSIS OF CONFUSION MATRIX

Figure 8 shows the confusion matrix obtained using MCnet [7]. The first row of this figure shows the confusion matrix in the hole ratio = 12.5%, the second row shows the confusion matrix in the hole ratio = 25%, and the third row shows the confusion matrix in the hole ratio = 50%. The first column represents the results obtained from the corrupted signal without any inpainting, the second column represents the results obtained from the signal restored by

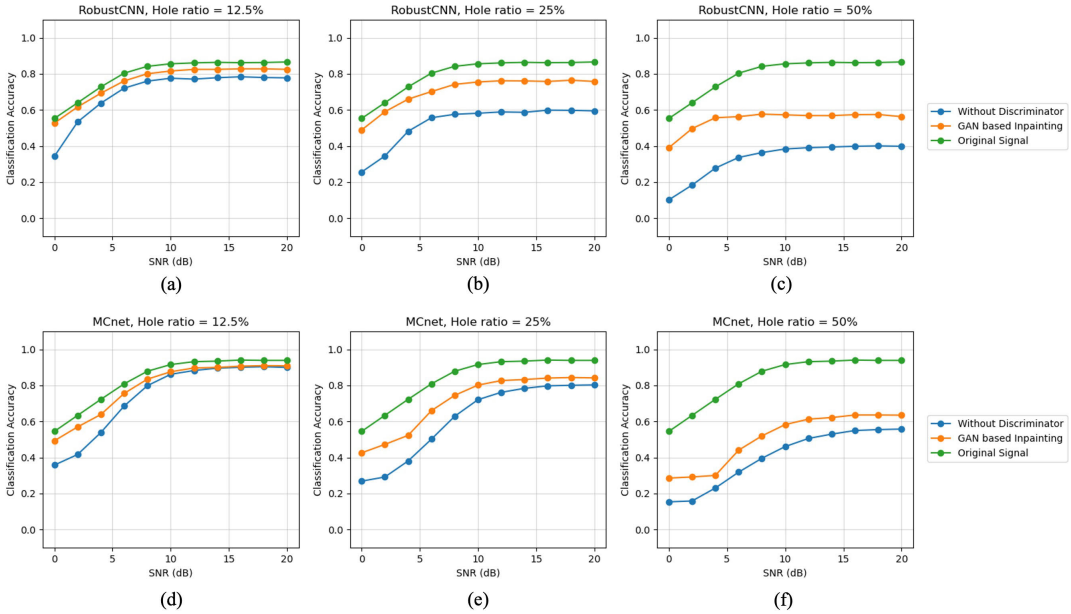


FIGURE 9. Effect of the discriminator. Each graph shows two variants trained with and without the discriminator in the proposed inpainting network.

the linear interpolation, and the third column represents the results obtained from the signal restored by the proposed GAN-based inpainting method.

In the case of the hole ratio = 12.5%, there is a tendency to misclassify high-order modulations when using the corrupted signal without inpainting. In particular, AM-SSB-WC and AM-SSB-SC are misclassified as 8-ASK. On the other hand, the results of the linear interpolation and the proposed inpainting method are relatively accurate in classifying AM-SSB-WC and AM-SSB-SC.

In addition, although the GAN-based inpainting shows more confusion for high-order modulations (APSK and QAM) and AM-SSB modulation categories compared to the confusion matrix of the original signal, it shows a similar tendency in the confusion matrix for most of the modulation categories.

In the case of the hole ratio = 25%, the overall tendency to confusion becomes even greater. In the linear interpolation, AM-DSB-WC modulation is misclassified as AM-DSB-SC modulation. In addition, some cases are observed where QPSK, 8-PSK, and FM signals are confused with 16-QAM. In contrast, the proposed inpainting method shows relatively accurate classification in these confusions.

In the case of the hole ratio = 50%, the corrupted signal shows inaccurate classification for most modulation types except for OOK. There is a tendency to misclassify 64-APSK, 128-APSK, 32-QAM, and AM-DSB-WC as 32-APSK. Also, in the case of the linear interpolation, although the confusion is reduced, there are still some misclassifications such as QAM modulation types.

In the case of the proposed inpainting method, there is still confusion about high-order modulations, compared to the original signal. For example, 32-APSK is misclassified as 128-APSK. Analog modulation types are also turned to be confused such as the confusion between AM-DSB-WC and AM-DSB-SC, and FM and 16-QAM. Nevertheless, the proposed inpainting model shows less confusion and higher overall accuracy.

Overall, this result indicates that, at the low hole ratio, the GAN-based inpainting shows similar performance on different modulation categories. However, at the high hole ratio, the inpainting have different performance on different modulation categories. For example, high-order modulations show lower performance due to their similarity in time-domain signals and large number of the missing samples to be inpainted. It should be noted that a further study is necessary to address the issues of such difficult and misclassifying modulation types.

5) EFFECT OF GAN ARCHITECTURE

In this section, the effectiveness of the GAN architecture in the modulation classification is analyzed through an ablation test. We compare the proposed GAN structure, which includes both the generator and discriminator, to a variant model trained with only the generator and L1 loss. The results of this ablation test are presented in Figure 9. In the figure, the first row is the result of the experiment with RobustCNN, and the second row is the result of the experiment with MCnet [7].

With the RobustCNN [8], the use of the discriminator improves an average performance of 6% and 18% at the hole ratios 12.5% and 25%, respectively. Also, at the 50% hole ratio, it achieves an average performance improvement of 21% in the modulation classification. Similarly, with the MCnet [7], the GAN structure improves the accuracy performance by an average of 5% in the hole ratio of 12.5%,

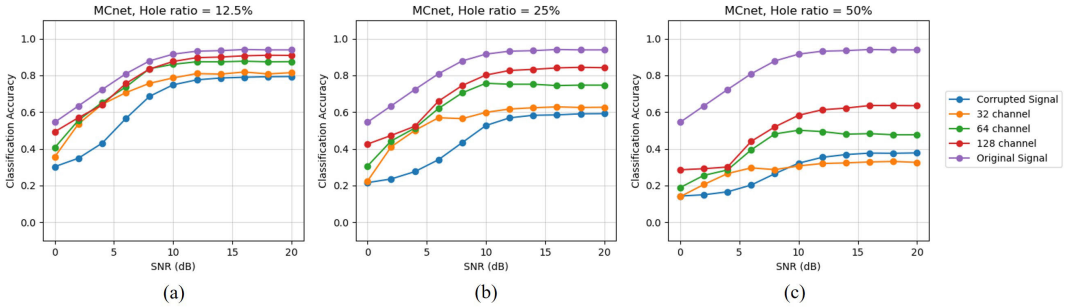


FIGURE 10. Ablation study for the number of channels in the first convolutional layer of the generator network. Note that three different numbers of the channel are used in this test.

compared to the model without the discriminator. At the hole ratios of 25% and 50%, the GAN structure shows an average performance improvement of 10%.

As the hole ratio increases, the performance improvement also increases. This is because more signals are generated as the number of holes increases, and these signals more resemble the data distribution of the original modulated signals due to the role of the discriminator. These experimental results demonstrate the effectiveness of the GAN network in IQ signal inpainting.

It should be further noted that, in our generator network, the first convolutional layer increases the signal channel from 1 (input) to 128. An ablation study was performed to decide to choose the number of channels, because finding an appropriate parameter is important to avoid overfitting and underfitting problem during the model training. To perform this ablation study, we used three different numbers of the channel for the first convolutional layer (i.e., 32, 64, and 128). Figure 10 shows the result of this ablation study. As shown in Figure 10, in case of 32 channels, the AMC accuracy was not much improved compared to that of the corrupted signal, even though it is slightly improved for low SNR signals. In case of 64 channels, the accuracy showed better improvement than that of 32 channels. However, the accuracy drops as the hole ratio increases. The best accuracy was obtained using 128 channels, as shown in Figure 10. Based on this ablation study, we selected 128 channels for the first convolution layer.

6) EFFECT OF SHORT LENGTH SIGNALS

The proposed inpainting model is a fully convolutional network, which means that it can work regardless of the length of an input signal. An additional experiment has been conducted to apply the proposed method to shorter length signals with 512 samples in the RadioML 2018.01A dataset [14].

Specifically, we made the signals of 512 samples by a random cropping of original 1024 samples from the RadioML 2018.01A dataset [14]. Due to the fully convolutional network, the inpainting model trained with 1024 signals was directly used to restore 512 signals without retraining the model. However, the backbone AMC model was retrained with the 512 signals and then tested with the 512 signals. Note that we used the RadioML 2018.01A dataset [14] to keep consistency between the inpainting and AMC models. Figure 11 shows the result of this experiment. As shown in Figure 11, the result shows a similar tendency between 1024 signals and 512 signals. This result indicates that the proposed inpainting model can be applied to shorter lengths and can be effective for the automatic modulation classification.

C. DISCUSSION ON REAL-WORLD IMPAIRED SIGNALS

In this section, we discuss the meaning and limitations of our study, when we apply the proposed method to the impaired signals in a real-world wireless transmission. The missing time samples in a real-world transmission are not modeled by zeros due to AWGN, attenuated signal, interference, etc. Hence, the missing samples should be modeled as real-world impaired signals rather than just zeros.

To investigate this practicality of the proposed method in a real-world transmission, we have performed an additional experiment to consider modeling the training data more realistically instead of just using zeros. Specifically, in this experiment, we replaced the missing samples with AWGN, while training the proposed GAN inpainting model. We used the same method to add Gaussian noise as mentioned in the specification of the RadioML 2018.01A dataset [14]. That is, we set the zero mean and the standard deviation of Gaussian noise as $std = 10^{-SNR/10}$. This means that lower SNR is, more noise is added. As a result of this experiment, we compared the AMC accuracy between the model with zeros (hereafter zero model) and the model with AWGN (hereafter AWGN model).

Figure 12 shows the comparison result of this experiment. As shown in this figure, at hole ratio = 12.5%, the AMC accuray shows the similar tendency between the zero model and AWGN model. At hole ratios = 25% and 50%, the same tendency is observed in case of above 10dB SNR, and the accuracy of the AWGN model is even higher than that of the zero model for low SNR signals. Note that, in case of high SNR signals, the noise has values almost close to zero. Hence, although the accuracy of the AWGN model is slightly decreased for above 10dB SNR, it is almost similar to that of the zero model. The classification accuracy of the AWGN

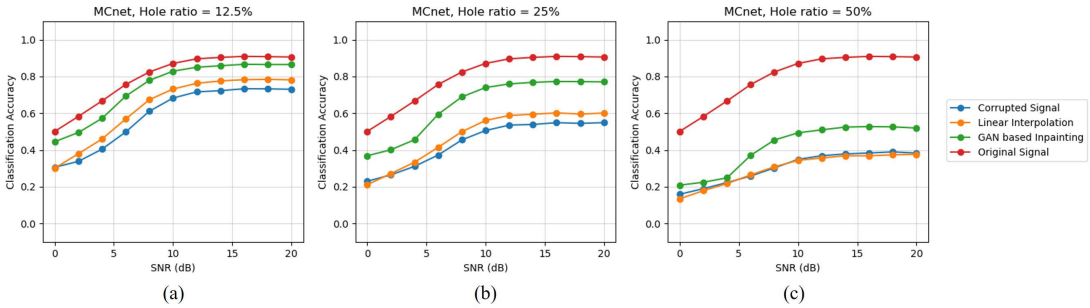


FIGURE 11. Effect of short length signal. In each graph, the GAN-based inpainting result was obtained from the signals with 512 samples.

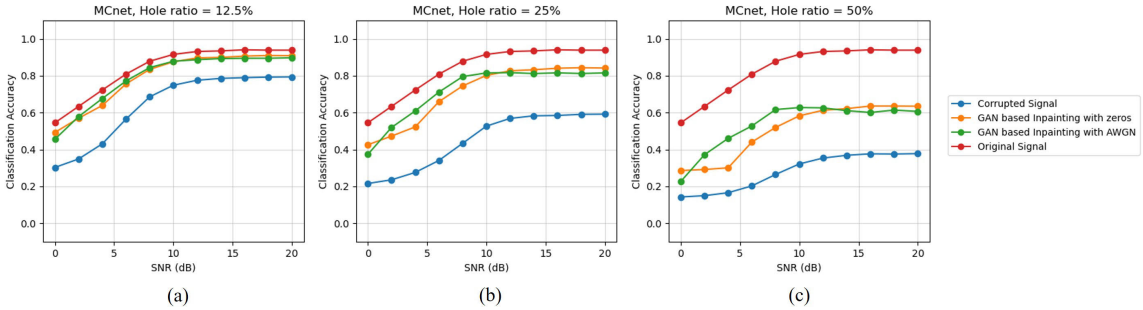


FIGURE 12. AMC accuracy of two inpainting models trained with the missing samples (one is modeled as zeros and another is modeled as AWGN). Each graph shows the accuracy result obtained by three different hole ratios (i.e., the percentage of missing samples in a signal).

model shows a maximum 8% higher than that of the zero model at hole ratio = 25%, and it shows a maximum 15% higher than that of the zero model at hole ratio = 50%. This is because the original (uncorrupted) low SNR signal itself contains many noises so that the AMC model can work better for the signals with noise. The result of this experiment indicates that the proposed method is capable of improving the AMC accuracy for more realistic data with AWGN.

Also, note that the RadioML 2018.01A dataset [14] does not provide the original (uncorrupted) transmission signal. Hence, it is difficult to regenerate/recover the missing samples so as to be modeled in the same conditions as the remaining neighbor samples. By this reason, we have considered missing samples with AWGN in this experiment. It should be further noted that although the proposed method is the first work to investigate the effect of deep learning-based signal inpainting on the automatic modulation classification, it remains as a future work to deploy the proposed method for the real-world impaired signals that considers all the channel conditions such as attenuated signal, interference, etc.

V. CONCLUSION

This study proposes a GAN-based signal inpainting method for restoring corrupted time-domain signals with missing samples in wireless communications. Also, intensive experiments were conducted to investigate the effect of signal inpainting on the modulation classification. The experimental results were obtained with RadioML dataset [14] used in the existing studies on AMC. Throughout the visual comparison and ablation study, the results indicate that the proposed signal inpainting method can generate the restored signal that resembles the characteristics of real signals. Particularly, it maintains the global structure of a time-domain signal for each modulation type, which enables the extraction of distinctive features for the modulation classification. More importantly, the GAN-based signal inpainting method improves the classification accuracy for most modulation types. In future works, it is necessary to verify the usefulness of the proposed method using a real dataset obtained in a real wireless communication environment, rather than the synthetic dataset obtained through a software simulation. Additionally, to be used in real communication systems, a low running time and computational complexity are required. Optimization of the network through lightweight techniques is necessary, and these tasks remain as important future works.

REFERENCES

[1] M. Amjad, M. H. Rehmani, and S. Mao, "Wireless multimedia cognitive radio networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1056–1103, 2nd Quart., 2018.

[2] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proc. Eng. Appl. Neural Netw., 17th Int. Conf. (EANN)*. Aberdeen, U.K.: Springer, Sep. 2016, pp. 213–226.

[3] J. L. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Trans. Syst., Man, Cybern., C Applications Reviews*, vol. 41, no. 4, pp. 455–469, Jul. 2011.

[4] D. Grimaldi, S. Rapuano, and L. De Vito, "An automatic digital modulation classifier for measurement on telecommunication networks," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 5, pp. 1711–1720, Oct. 2007.

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

[6] 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.

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

[8] S. Kim, J. Kim, W. Nwadiugwu, and D. Kim, "Deep learning-based robust automatic modulation classification for cognitive radio networks," *IEEE Access*, vol. 9, pp. 92386–92393, 2021.

[9] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3072–3108, 4th Quart., 2019.

[10] Q. Zheng, P. Zhao, Y. Li, H. Wang, and Y. Yang, "Spectrum interference-based two-level data augmentation method in deep learning for automatic modulation classification," *Neural Comput. Appl.*, vol. 33, no. 13, pp. 7723–7745, Jul. 2021.

[11] T. Huynh-The, Q. Pham, T. Nguyen, T. T. Nguyen, R. Ruby, M. Zeng, and D. Kim, "Automatic modulation classification: A deep architecture survey," *IEEE Access*, vol. 9, pp. 142950–142971, 2021.

[12] E. Blossom, "Gnu radio: Tools for exploring the radio frequency spectrum," *Linux J.*, vol. 2004, no. 122, p. 4, 2004.

[13] T. J. O'shea and N. West, "Radio machine learning dataset generation with GNU radio," in *Proc. GNU Radio Conf.*, vol. 1, 2016, pp. 1–6.

[14] 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.

[15] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, "Image inpainting for irregular holes using partial convolutions," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 85–100.

[16] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. Huang, "Free-form image inpainting with gated convolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4471–4480.

[17] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5505–5514.

[18] S. M. N. Uddin and Y. J. Jung, "Global and local attention-based free-form image inpainting," *Sensors*, vol. 20, no. 11, p. 3204, Jun. 2020.

[19] E. Ntavelis, A. Romero, S. Bigdeli, R. Timofte, Z. Hui, X. Wang, X. Gao, C. Shin, T. Kim, H. Son, and S. Lee, "AIM 2020 challenge on image extreme inpainting," in *Computer Vision—ECCV*. Glasgow, U.K.: Springer, Aug. 2020, pp. 716–741.

[20] S. M. N. Uddin and Y. J. Jung, "SIFNet: Free-form image inpainting using color split-inpaint-fuse approach," *Comput. Vis. Image Understand.*, vol. 221, Aug. 2022, Art. no. 103446.

[21] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[22] F. Restuccia, S. D'Oro, A. Al-Shawabka, B. Costa Rendon, K. Chowdhury, S. Ioannidis, and T. Melodia, "Hacking the waveform: Generalized wireless adversarial deep learning," 2020, *arXiv:2005.02270*.

[23] T. Roy, T. O'Shea, and N. West, "Generative adversarial radio spectrum networks," in *Proc. ACM Workshop Wireless Secur. Mach. Learn.*, May 2019, pp. 12–15.

[24] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks," *IEEE Access*, vol. 6, pp. 15713–15722, 2018.

[25] A. Marafioti, N. Perraudin, N. Holighaus, and P. Majdak, "A context encoder for audio inpainting," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 27, no. 12, pp. 2362–2372, Dec. 2019.

[26] Y. Bahat, Y. Y. Schechner, and M. Elad, "Self-content-based audio inpainting," *Signal Process.*, vol. 111, pp. 61–72, Jun. 2015.

[27] N. E. Barabanov and D. V. Prokhorov, "Stability analysis of discrete-time recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 292–303, Mar. 2002.

[28] Q. Song, Y. Wu, and Y. C. Soh, "Robust adaptive gradient-descent training algorithm for recurrent neural networks in discrete time domain," *IEEE Trans. Neural Netw.*, vol. 19, no. 11, pp. 1841–1853, Nov. 2008.

[29] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, Sep. 2018.

[30] S. Chandhok, H. Joshi, S. J. Darak, and A. V. Subramanyam, "LSTM guided modulation classification and experimental validation for sub-nyquist rate wideband spectrum sensing," in *Proc. 11th Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2019, pp. 458–460.

[31] Y. Chen, W. Shao, J. Liu, L. Yu, and Z. Qian, "Automatic modulation classification scheme based on LSTM with random erasing and attention mechanism," *IEEE Access*, vol. 8, pp. 154290–154300, 2020.

[32] L. Huang, W. Pan, Y. Zhang, L. Qian, N. Gao, and Y. Wu, "Data augmentation for deep learning-based radio modulation classification," *IEEE Access*, vol. 8, pp. 1498–1506, 2020.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

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

[35] T. Huynh-The, V. Doan, C. Hua, Q. Pham, and D. Kim, "Chain-Net: Learning deep model for modulation classification under synthetic channel impairment," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2020, pp. 1–6.

[36] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, and Y. 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.

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[38] 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)*, Jun. 2015, pp. 1–9.

[39] 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.

[40] 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.

[41] J. Cai, F. Gan, X. Cao, and W. Liu, "Signal modulation classification based on the transformer network," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 3, pp. 1348–1357, Sep. 2022.

[42] S. Hamidi-Rad and S. Jain, "MCformer: A transformer based deep neural network for automatic modulation classification," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

[43] Q. Zheng, P. Zhao, H. Wang, A. Elhanashi, and S. Saponara, "Fine-grained modulation classification using multi-scale radio transformer with dual-channel representation," *IEEE Commun. Lett.*, vol. 26, no. 6, pp. 1298–1302, Jun. 2022.

[44] P. Smaragdis, B. Raj, and M. Shashanka, "Missing data imputation for spectral audio signals," in *Proc. IEEE Int. Workshop Mach. Learn. Signal Process.*, Sep. 2009, pp. 1–6.

[45] P. J. Wolfe and S. J. Godsill, "Interpolation of missing data values for audio signal restoration using a Gabor regression model," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2005, pp. 1–4.

[46] P. P. Ebner and A. Eltelt, "Audio inpainting with generative adversarial network," 2020, *arXiv:2003.07704*.

[47] Y.-L. Chang, K.-Y. Lee, P.-Y. Wu, H.-Y. Lee, and W. Hsu, "Deep long audio inpainting," 2019, *arXiv:1911.06476*.

[48] H. Zhou, Z. Liu, X. Xu, P. Luo, and X. Wang, "Vision-infused deep audio inpainting," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 283–292.

[49] M. Sadeghi and E. G. Larsson, "Adversarial attacks on deep-learning based radio signal classification," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 213–216, Feb. 2019.

[50] R. Sahay, C. G. Brinton, and D. J. Love, "A deep ensemble-based wireless receiver architecture for mitigating adversarial attacks in automatic modulation classification," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 1, pp. 71–85, Mar. 2022.

[51] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.-Assist. Intervent., 18th Int. Conf. (MICCAI)*, Munich, Germany: Springer, Oct. 2015, pp. 234–241.
[52] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, *arXiv:1511.07122*.
[53] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
[54] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.
[55] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*.
[56] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, and L. Antiga, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, 32, 2019, pp. 1–12.
[57] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

SUBIN LEE received the B.S. degree from the Department of Software, Gachon University, South Korea, in 2022. Since 2021, he has been a Researcher with the Computer Vision and Image Processing Laboratory (CVIP Lab), Gachon University. His current research interests include computer vision and deep learning.

YOUNG-IL YOON received the M.S. degree from Chungnam National University, Daejeon, South Korea, in 2013.

From 2012 to 2013, he was a Software Engineer with the Cloud Team, Naver Cloud, contributing to the virtual machine. Since 2013, he has been the Project Manager with the C4I Research Center, LIG Nex1, contributing to the tactical radio systems. His current research interests include software frameworks and embedded Linux kernel.

YONG JU JUNG (Member, IEEE) received the Ph.D. degree from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2005.

From 2005 to 2010, he was a Principal Research Scientist with the Samsung Advanced Institute of Technology, contributing to 3D display processing for 3D TV. From 2010 to 2014, he was an Associate Research Professor with the Image and Video Systems Laboratory, Department of Electrical Engineering, KAIST. From 2014 to 2016, he was a Principal Engineer with the System LSI Division, Samsung Electronics, contributing to the development of image sensors and multi-camera solutions. Since 2016, he has been an Associate Professor with the School of Computing and the Director of Computer Vision and Image Processing Laboratory (CVIP Lab), Gachon University. His current research interests include signal processing, computer vision, and deep learning. He was a recipient of the Samsung Muhan Research Award in 2010 and the Best Innovation Award from Samsung Electronics in 2015. He co-organized special sessions on human 3D perception and 3D video assessments in DSP2011. He is an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.