

Deep Learning-based ISI and ICI Suppression in OTFS-based ISAC Systems

Soyoon Park and Seongwook Lee
Department of Electrical and Electronics Engineering
College of ICT Engineering, Chung-Ang University
Seoul, Republic of Korea
{thdbs0711, seongwooklee}@cau.ac.kr

Abstract—This paper proposes a deep learning-based method for suppressing inter-symbol interference (ISI) and inter-carrier interference (ICI) in orthogonal time frequency space (OTFS)-based systems. In conventional OTFS-based systems, the accuracy of channel estimation degrades due to ISI and ICI. To address this problem, we design a neural network that restores the time-domain received signal by removing the effects of ISI and ICI while preserving the channel characteristics. The proposed network adopts the U-Net architecture with attention blocks, and it is trained on simulated datasets that cover diverse numbers of targets and a wide range of signal-to-noise ratio values. We evaluate the training process based on the measure of loss convergence to ensure stable learning. The performance improvement is then quantified by the root mean square error (RMSE), which is reduced from 0.42 to 0.30. Finally, the effectiveness of the proposed method is verified by reconstructing the target profile in the OTFS-based sensing system, showing successful restoration even at 0 dB. These findings demonstrate that the proposed method effectively mitigates ISI and ICI and provides robust performance of target detection in OTFS-based sensing systems.

Index Terms—deep learning, inter-carrier interference (ICI), inter-symbol interference (ISI), orthogonal time frequency space (OTFS).

I. INTRODUCTION

Modern wireless communication systems are required to maintain reliable performance while achieving high spectral efficiency. In sixth-generation communications, integrated sensing and communication (ISAC) is emerging as a core technology that provides both stable communication and precise sensing in high-mobility environments [1], [2]. Orthogonal time frequency space (OTFS) has attracted significant attention as a modulation technique for next-generation ISAC systems because it can simultaneously support communication and sensing in high-mobility environments [3]. OTFS assigns data symbols in the delay-Doppler (DD) domain, rather than assigning them in the conventional time-frequency (TF) domain. The DD domain directly reflects the fundamental channel characteristics of delay and Doppler shift, and it remains quasi-static even in channels with high Doppler shifts [4], [5]. As a result, OTFS offers the advantage of maintaining reliable performance under high mobility and severe fading conditions compared to TF domain-based modulation schemes [6], [7].

In practical channel conditions, OTFS-based systems are affected by several factors of performance degradation. The

most prominent among these are inter-symbol interference (ISI) and inter-carrier interference (ICI). When ISI occurs, the delay spread of the channel causes signal overlap between adjacent symbols, and its impact becomes more severe as the spread increases [8], [9]. Similarly, when ICI occurs, Doppler spread induces interference between adjacent Doppler bins. The effect intensifies with increasing Doppler shift induced by the channel [10], [11].

Representative conventional methods to mitigate such interference are cyclic prefix (CP) and zero padding (ZP). The CP is generated by copying the end of the symbol to its beginning so that linear convolution is converted into circular convolution. The addition of CP reduces the data rate per unit time under a fixed bandwidth and therefore decreases spectral efficiency [12]. In the case of ZP, zero values are inserted at the beginning of a symbol to eliminate symbol overlap [13]. Under the same transmit power constraint, the inclusion of a zero interval reduces the average power allocated to the valid symbols, which results in a loss of signal-to-noise ratio (SNR). In addition, this also leads to an increase in the peak-to-average power ratio (PAPR), which is undesirable for practical systems. Furthermore, the lengths of CP and ZP must exceed the maximum channel delay, which can lead to significant overhead when channels with large delay spreads are considered. These conventional methods suffer limitations such as spectral efficiency reduction and power loss. To overcome them, we propose a deep learning-based approach for ISI and ICI suppression. The proposed method combines an attention block with a U-Net-based network to accurately predict and remove complex interference patterns. By directly learning and suppressing interference components in the time-domain signal, the proposed method restores an interference-free signal, thereby significantly enhancing the performance of OTFS-based sensing systems.

II. SIGNAL MODEL OF OTFS-BASED SYSTEMS

In the OTFS modulation scheme, the information bits are assigned to complex-valued quadrature amplitude modulation (QAM) symbols. The modulated QAM symbols are mapped onto the DD domain, where the symbol located at Doppler bin k and delay bin l is denoted by $x_{DD}[k, l]$. These symbol representations in the DD domain are subsequently transformed into the TF domain via the inverse symplectic fast Fourier

transform. The transmitted signal in the TF domain can be expressed as

$$x_{\text{TF}}[k', m] = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} x_{\text{DD}}[k, l] \exp \left(j2\pi \left(\frac{ml}{M} - \frac{kk'}{N} \right) \right), \quad (1)$$

where M denotes the number of subcarriers, N denotes the number of time slots, and k' and m denote the subcarrier and time bins in the TF domain, respectively. The TF domain signal is converted into the time domain by applying the Heisenberg transform, and the resulting time-domain signal is upconverted and then transmitted, which can be expressed as

$$\begin{aligned} x(t) &= \sum_{k'=0}^{N-1} \sum_{m=0}^{M-1} x_{\text{TF}}[k', m] \\ &\times \exp(j2\pi m(f_c + k'\Delta f)(t - mT)) \\ &\times \text{rect}\left(\frac{t - mT}{T}\right). \end{aligned} \quad (2)$$

Here, f_c denotes the carrier frequency, Δf is the subcarrier spacing, and T is the symbol duration. After propagating through the channel, the transmitted signal is received and downconverted to baseband for subsequent processing. The received signal is transformed back to the DD domain by applying the symplectic fast Fourier transform and the Wigner transform. The resulting DD-domain received signal can be approximated as a two-dimensional (2D) convolution between the channel response and the transmitted signal, expressed as

$$y_{\text{DD}}[k, l] \approx \sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} h[k', l'] x_{\text{DD}}[[k - k']_N, [l - l']_M], \quad (3)$$

where $h[k', l']$ denotes the channel response in the DD domain. Here, $[\cdot]_N$ and $[\cdot]_M$ represent the modulo- N and modulo- M operations, respectively.

In the OTFS-based sensing system, the channel response $h[k', l']$ in (3) can be interpreted as the target response, which includes information about the range and velocity of the target. Then, the range and velocity of the target can be obtained through 2D deconvolution, as follows: the received signal in the DD domain is first converted to the frequency domain via the 2D discrete Fourier transform. Spectrum separation is then performed to eliminate the influence of data symbols and to retain only the components associated with the target. Subsequently, the 2D inverse discrete Fourier transform is applied to transform the signal back into the DD domain. Finally, the range and velocity of the target can be directly estimated from the processed DD-domain signal.

III. PERFORMANCE DEGRADATION DUE TO ISI AND ICI IN OTFS-BASED SYSTEMS

In the ideal case, the received signal in the DD domain can be approximated as a 2D convolution between the channel response and the transmitted signal as in (3). However, in practical systems the received signal in the DD domain cannot be expressed in the form of (3), because delay and Doppler

spread of the channel result in ISI and ICI, respectively. Specifically, ISI arises when the maximum delay spread of the channel exceeds the guard interval, such as CP and ZP. Similarly, ICI results from Doppler spread of the channel, which leads to coupling between adjacent Doppler bins. As a result, additional phase terms are multiplied with the received signal, and the received signal can no longer be approximated as a simple 2D convolution. The DD-domain received signal that accounts for these effects can be expressed as

$$\begin{aligned} y_{\text{DD}}[k, l] &= \sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} h[k', l'] x_{\text{DD}}[[k - k']_N, [l - l']_M] \\ &\times \exp \left(j \frac{2\pi}{MN} [k - k']_N l' \right) \\ &\times \exp \left(j \frac{2\pi}{M} l u(-k + k') \right), \end{aligned} \quad (4)$$

where the first and second exponential terms are due to ICI and ISI, respectively, and $u(\cdot)$ denotes the unit step function.

In the presence of ISI and ICI, the reflected signal energy is not concentrated in the DD grid point that represents the channel-induced delay and Doppler shift. Instead, it disperses into adjacent bins, which causes spectral leakage. This dispersion lowers the effective SNR of the corresponding DD grid point. In the OTFS-based sensing system, the target response usually concentrates within a single DD grid point. Detection performance can be significantly degraded when the attenuated response is further corrupted by noise. Furthermore, when this attenuation is added to noise, the target may become undetectable, which can be observed in Fig. 1 (a). As shown in Fig. 1 (a), when ISI and ICI are taken into account, only Target 1 with a small delay bin is clearly detected. In contrast, Fig. 1 (b) shows that when ISI and ICI are absent, both Target 1 and Target 2 are successfully detected. This implies that OTFS is inherently robust against ICI caused by Doppler spread but vulnerable to ISI. Such an observation can also be explained by the structure of the phase terms. The phase term corresponding to ICI can be expressed as $\exp(j \frac{2\pi}{MN} [k - k']_N l')$, which is scaled by the frame size MN . Therefore, when MN is sufficiently large, the accumulated interference is naturally suppressed. On the other hand, the phase term corresponding to ISI can be expressed as $\exp(j \frac{2\pi}{M} l u(-k + k'))$. As the signal delay increases due to the channel, phase dispersion becomes more severe, and the energy spreads into adjacent delay bins. Therefore, in the OTFS-based system, performance degradation is mainly caused by ISI, which critically limits the detection of distant targets.

IV. PROPOSED METHOD FOR INTERFERENCE MITIGATION

A. Configuration of Training Dataset

The objective of the proposed method is to reconstruct the time-domain received signal by removing the effects of ISI and ICI, thereby approximating an interference-free signal. To achieve this goal, a deep learning-based suppression approach for ISI and ICI is suggested. The neural network is trained

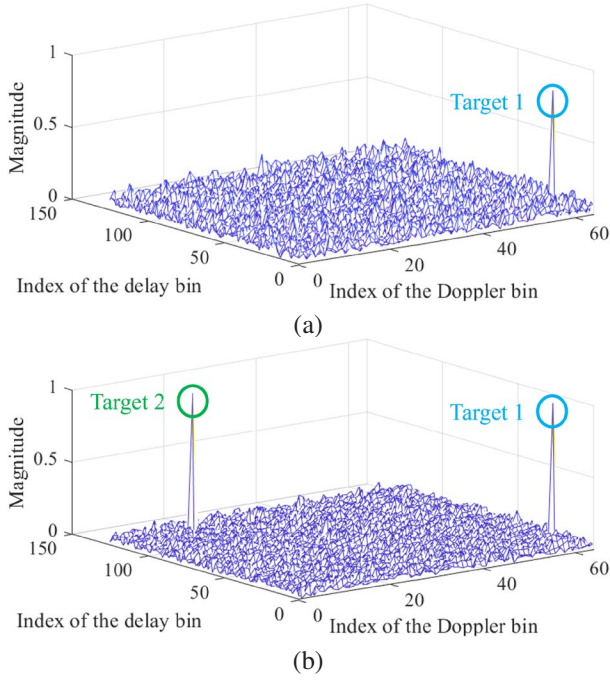


Fig. 1. Comparison of target responses in the DD domain for Target 1 (delay bin: 20, Doppler bin: 61) and Target 2 (delay bin: 118, Doppler bin: 15): (a) with ISI and ICI effects and (b) without ISI and ICI effects.

with the time-domain received signal affected by ISI and ICI as the input, and the corresponding interference-free signal as the ground truth. In this way, the network learns to mitigate ISI and ICI. The time-domain signal not only reflects the influence of ISI and ICI but also retains the complete channel characteristics. These features allow the network to effectively capture and suppress interference patterns. In addition, once the restoration is performed in the time-domain, subsequent signal processing can be performed based on the ideal signal, which extends the applicability of the proposed method. The training dataset is generated through simulations that reflect diverse scenarios, and the system parameters adopted in this paper are summarized in Table I. Specifically, both single-target and multi-target cases are considered by setting the number of targets to range from 1 to 5. Furthermore, the SNR is varied from 0 dB to 30 dB so that the network can be trained under various SNR values.

TABLE I
PARAMETERS USED IN SIMULATIONS

Parameter	Value
Data modulation scheme	16-QAM
Center frequency (GHz)	24
Bandwidth (MHz)	61.44
The number of Doppler bins	64
The number of delay bins	128

B. Structure of Proposed Neural Network

In this paper, we propose the deep learning-based model that removes ISI and ICI from the time-domain signal. It restores clean signals by applying the U-Net architecture together with attention blocks. The structure of the proposed neural network is shown in Fig. 2. The input stage of the network receives two channels corresponding to the in-phase and quadrature components of the time-domain received signal. These inputs are converted by a convolution (Conv) layer into feature maps suitable for interference mitigation. The resulting feature maps are also stored as an input residual, which is used in the final output stage.

The encoder performs hierarchical feature extraction through 4 stages, where each stage is composed of a Conv block and a max pooling operation. Within each Conv block, a Conv operation is followed by batch normalization (BN) and a rectified linear unit (ReLU) activation function. BN normalizes the distribution of feature maps across channels to enhance training stability. In turn, ReLU offers nonlinearity, allowing the network to capture complex interference patterns. After this process, dropout is applied to prevent overfitting, and the second Conv operation is performed. The output is subsequently passed through BN and ReLU once again, which enhances the ability to extract features of the network. In the first block, the number of input channels is expanded from 2 to 32, and in the subsequent blocks the number of channels is increased step by step to 64, 128, and 256. Finally, max pooling gradually reduces the spatial resolution, which facilitates the learning of global patterns. The final output of the encoder is expressed as a feature map with 512 channels obtained through a Conv layer in the bottleneck stage, and this is followed by a residual block. The residual block contains 2 Conv layers together with BN and ReLU, and it is designed such that the input tensor is added directly before it enters the activation function. This structure alleviates the vanishing gradient problem that may occur in deep networks and improves training efficiency.

The decoder restores spatial resolution through 4 stages of upsampling, and each stage performs a skip connection using an attention block. In the attention block, the upsampled feature map from the decoder and the corresponding feature map from the encoder are each passed through a Conv layer to extract important features from both inputs. The resulting feature maps are then added elementwise, and the output is sequentially processed by ReLU, Conv, BN, and finally a sigmoid function to generate an attention mask. This mask is multiplied with the encoder feature map so that only the information at important spatial locations is emphasized. The feature map of the encoder enhanced by the attention block captures fine local details. In contrast, the feature map of the decoder after upsampling represents global contextual information, and together they provide complementary information. The masked feature map of the encoder, obtained after sigmoid activation, is concatenated with the upsampled feature map of the decoder along the channel dimension, which results

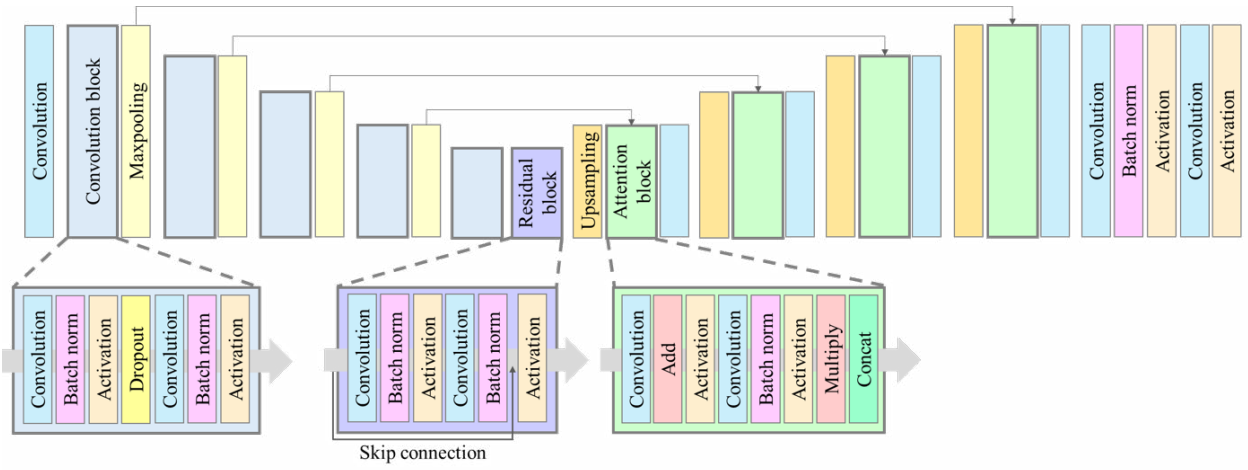


Fig. 2. Architecture of the proposed neural network for ISI and ICI mitigation in the OTFS-based system.

in a doubling of the number of channels. The Conv layer is subsequently applied to fuse the local and global features while reducing the number of channels. The attention mechanism directs the decoder to encoder features that are most relevant for reconstructing interference components. Less informative regions are suppressed, which helps the network produce more accurate interference predictions.

The final output of the decoder is transformed into the prediction of the interference through the output module. The output module sequentially reduces the number of channels to 16 and then to 2 through each Conv layer, and applies the hyperbolic tangent (Tanh) as the activation function. The Tanh constrains the output range, which enhances training stability and prevents the network from overestimating the interference magnitude. The clean signal without interference is finally obtained by subtracting the predicted interference signal from the stored input residual. This architecture encourages the network to predict the interference components instead of directly generating the restored signal, which simplifies the learning process and accelerates convergence.

C. Loss Function

To train the proposed network, we adopt a loss function that evaluates the prediction errors of both the real and imaginary components of the received signal. Specifically, the mean squared error (MSE) is computed for each channel independently, and the two errors are aggregated to form the final loss. The loss function is defined as

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N \left((y_{\text{Re},i} - \hat{y}_{\text{Re},i})^2 + (y_{\text{Im},i} - \hat{y}_{\text{Im},i})^2 \right), \quad (5)$$

where $y_{\text{Re},i}$ and $\hat{y}_{\text{Re},i}$ denote the ground truth and predicted values of the real channel, and $y_{\text{Im},i}$ and $\hat{y}_{\text{Im},i}$ denote those of the imaginary channel, respectively. Here, N represents the total number of samples.

D. Setting of Training Parameters

The experiments were implemented using the PyTorch framework. The training dataset consisted of 5000 samples, with 80 percent allocated for training and 20 percent reserved for validation. Because signals from distant targets are more vulnerable to ISI, the dataset was designed to emphasize such cases. Specifically, when the total number of delay bins is 128, 70% of the samples were generated with target delays between 60 and 125. The network was trained for 50 epochs with a batch size of 32 using the adaptive moment estimation optimizer with a learning rate of 0.0002.

V. PERFORMANCE EVALUATION

The performance of the proposed method was evaluated through the convergence of the loss function, the root mean square error (RMSE) metric, and the restoration results of the target response. First, the value of the loss function during training is shown in Fig. 3. The loss values for both the training and validation datasets decreased steadily and converged to sufficiently small values after approximately 10 epochs. This confirms that the proposed network was trained stably and achieved generalization capability without overfitting. Also, the RMSE was calculated for quantitative evaluation. The proposed network reduced the RMSE from 0.42 to 0.30. This indicates that the proposed method effectively mitigates interference and produces outputs that are similar to the ground truth.

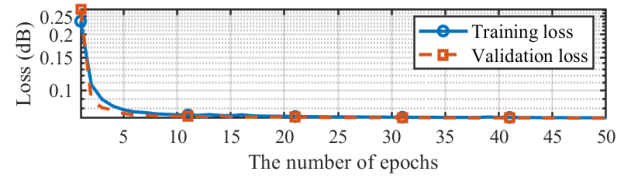


Fig. 3. Training and validation loss of the proposed network.

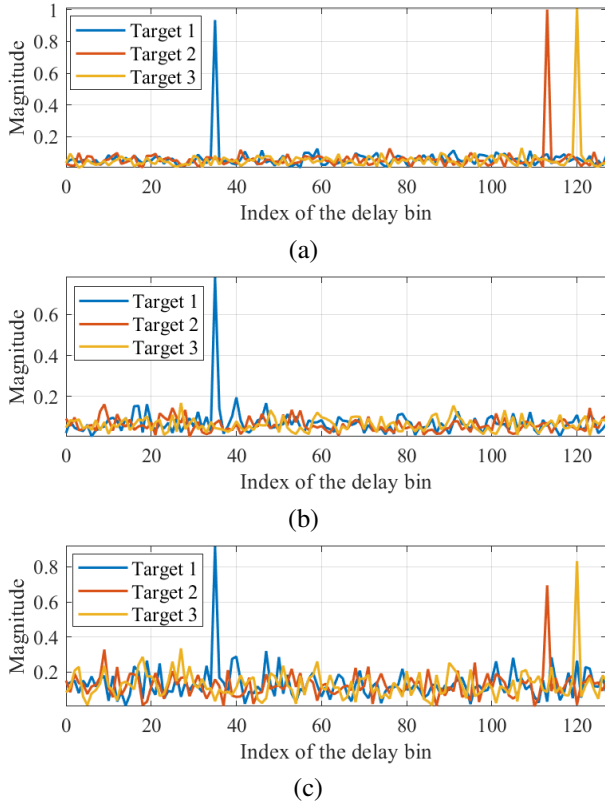


Fig. 4. The delay profiles: (a) ground truth signal, (b) input signal with ISI/ICI, and (c) restored output signal.

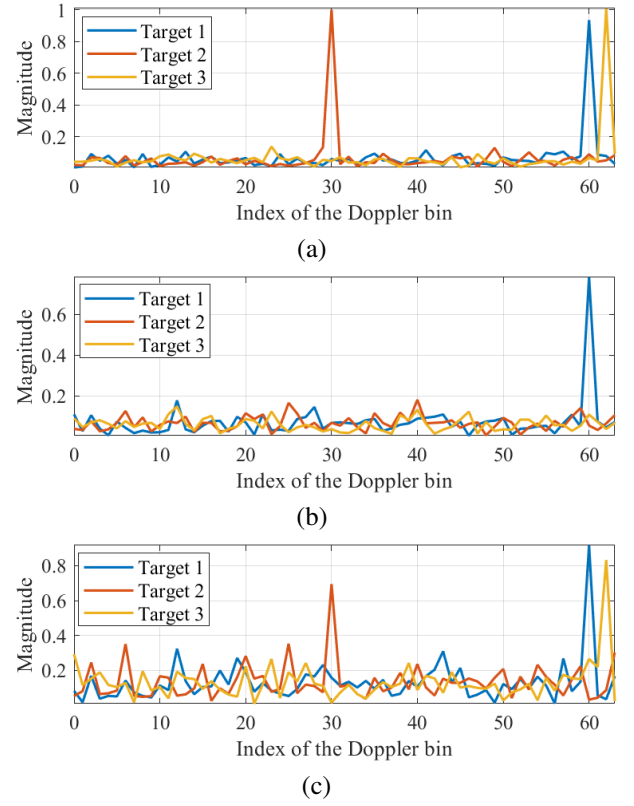


Fig. 5. The Doppler profiles: (a) ground truth signal, (b) input signal with ISI/ICI, and (c) restored output signal.

Finally, based on the received signals obtained from the proposed method, the sensing performance was evaluated through simulations focusing on reconstruction of the target response. For this evaluation, the received signals were generated at an SNR of 0 dB using arbitrary inputs that were not part of the training dataset. Fig. 4 illustrates the delay profiles of the target responses, and Fig. 5 shows the corresponding Doppler profiles. Each figure compares the ground truth signal, the input signal with interference, and the output signal reconstructed by the proposed network. The three targets are located at distinct positions in the DD grid: Target 1 with the 35th delay bin and the 60th Doppler bin, Target 2 with the 113th delay bin and the 30th Doppler bin, and Target 3 with the 120th delay bin and the 62nd Doppler bin.

In the input signal with interference, Target 2 and Target 3 observed at large delay bins were not properly detected. In contrast, Target 1, which has a relatively small delay bin but a high Doppler bin, was successfully detected. With the proposed method, the delay and Doppler profiles were successfully restored. This result shows that interference was effectively suppressed, and targets with large delays or high Doppler shifts were accurately restored without performance degradation.

VI. CONCLUSION

This paper proposed a deep learning-based method to suppress ISI and ICI in the OTFS-based system. The method

was designed to restore the time-domain received signal by training a U-Net-based architecture with attention blocks to predict and remove interference components. The training dataset was designed to reflect diverse environments under various SNR conditions and target scenarios. Each sample was constructed as a pair, with the ISI/ICI-distorted signal as the input and the corresponding interference-free signal as the ground truth. The proposed network converged stably during training and demonstrated strong generalization performance. Simulation results showed that the RMSE was reduced from 0.42 for the input signal to 0.30 for the output of the proposed method. Furthermore, at an SNR of 0 dB the proposed method successfully restored delay and Doppler profiles similar to those obtained from the ground truth signal. These results demonstrate that the proposed method effectively mitigates ISI and ICI in the OTFS-based system and provides robust target detection even in low-SNR values.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00405510).

REFERENCES

- [1] A. Kaushik, R. Singh, S. Dayarthna, R. Senanayake, M. D. Renzo, and M. Dajer, "Toward integrated sensing and communications for 6G: key enabling technologies, standardization, and challenges," *IEEE*

Communications Standards Magazine, vol. 8, no. 2, pp. 52–59, June 2024.

- [2] T. Jeong, C. Park, and S. Lee, “Geometric-sequence-decomposition-based joint range and velocity estimation in OFDM radar system for UAM applications,” *IEEE Internet of Things Journal*, vol. 11, no. 9, pp. 16941–16953, May 2024.
- [3] S. Li, W. Yuan, C. Liu, Z. Wei, J. Yuan, and B. Bai, “A novel ISAC transmission framework based on spatially-spread orthogonal time frequency space modulation,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1854–1872, June 2022.
- [4] H. Liu, Y. Liu, M. Yang, and Q. Zhang, “On the characterizations of OTFS modulation over multipath rapid fading channel,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 3, pp. 2008–2021, March 2023.
- [5] W. Yuan, S. Li, Z. Wei, Y. Cui, J. Jiang, and H. Zhang, “New delay Doppler communication paradigm in 6G era: a survey of orthogonal time frequency space (OTFS),” *China Communications*, vol. 20, no. 6, pp. 1–25, June 2023.
- [6] S. K. Mohammed, “Time-domain to delay-Doppler domain conversion of OTFS signals in very high mobility scenarios,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 6178–6183, June 2021.
- [7] L. Gaudio, G. Colavolpe, and G. Caire, “OTFS vs. OFDM in the presence of sparsity: a fair comparison,” *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 4410–4423, June 2022.
- [8] H. Qu, G. Liu, L. Zhang, S. Wen, and M. A. Imran, “Low-complexity symbol detection and interference cancellation for OTFS system,” *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1524–1537, March 2021.
- [9] M. F. Keskin, C. Marcus, O. Eriksson, A. Alvarado, J. Widmer, and H. Wymeersch, “Integrated sensing and communications with MIMO-OTFS: ISI/ICI exploitation and delay-Doppler multiplexing,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 8, pp. 10229–10246, August 2024.
- [10] A. Tusha and H. Arslan, “Low complex inter-Doppler interference mitigation for OTFS systems via global receiver windowing,” *IEEE Transactions on Vehicular Technology*, vol. 72, no. 6, pp. 7685–7698, June 2023.
- [11] Y. Ma, G. Ma, B. Ai, D. Fei, N. Wang, and Z. Zhong, “Characteristics of channel spreading function and performance of OTFS in high-speed railway,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 10, pp. 7038–7054, October 2023.
- [12] S. E. Zeggar and H. Arslan, “A novel cyclic prefix configuration for enhanced reliability and spectral efficiency in OTFS systems,” *IEEE Wireless Communications Letters*, vol. 12, no. 5, pp. 888–892, May 2023.
- [13] H. Qu, G. Liu, M. A. Imran, S. Wen, and L. Zhang, “Efficient channel equalization and symbol detection for MIMO OTFS systems,” *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 6672–6686, August 2022.