

BOC-FPDNet: BOC First arrival Path Detection in Multipath Channel for GNSS Positioning

Sangjae Cho, *Student Member, IEEE*, Dong-Hee Paek, *Student Member, IEEE*,
Hong-Woo Seok, *Student Member, IEEE*, and Seung-Hyun Kong*, *Senior Member, IEEE*

Abstract—Binary Offset Carrier (BOC) modulation is used in the Global Navigation Satellite System (GNSS). Unlike Binary Phase Shift Keying (BPSK) in the Global Positioning System, BOC-modulated signals have side-peaks in the autocorrelation function (ACF) leading to positioning errors in multipath channels. This paper introduces a two-stage BOC first arrival path detection network (BOC-FPDNet). This includes side-peaks cancellation (SC) and multipath time delay estimation (MDE) for BOC modulation. The SC is a deep learning-based method that cancels side-peaks without added computational load or noise increase. The MDE uses Multi-head attention to extract ACF output features and consists of two heads, positive and negative sample detector, for time delay estimation. The effectiveness of BOC-FPDNet is demonstrated by comparing it to super-resolution and other deep learning detection techniques in various multipath channel simulations and real urban data. Results show BOC-FPDNet provides superior side-peak cancellation and accurate BOC-modulated signal detection.

Index Terms—GNSS, BOC, Multipath.

I. INTRODUCTION

Today, positioning has become a key technology spanning various domains, from smartphones to autonomous vehicles. Outdoor positioning to date has heavily relied on Global Navigation Satellite Systems (GNSS) such as the Global Positioning System (GPS), Galileo, Glonass, and Beidou [1]. Many GNSSs employ Binary Phase Shift Keying (BPSK) modulation, a form of direct sequence spread spectrum (DSSS) where data is transmitted over a broad bandwidth due to system-specific pseudorandom noise (PN) codes. Other GNSSs have adopted a range of split-spectrum modulations known as Binary Offset Carrier (BOC) modulation, aiming to enhance positioning performance and to mitigate inter-system interference [2]–[5].

In the process of the positioning using satellite signals, GNSS receivers undergo two primary stages: Acquisition and Tracking. These stages correspond to coarse and fine searches, respectively. In both stages, the correlation between the received signal and the replica PN sequence generated by the receiver is analyzed to detect the signal's time delay. This time delay incorporates errors due to influences from the ionosphere, troposphere, and multipath channels during the satellite signal's propagation. Advances in technology have

ushered in continuous improvements in common error modeling and correction techniques related to the troposphere and ionosphere in GNSS [6], [7]. However, the multipath channels, where wireless signals encounter reflection, diffraction, and scattering due to surrounding obstacles and thus reach the receiver through various paths, are highly dependent on the surrounding environment. As a result, it's challenging to model it mathematically or to estimate its associated errors.

In the multipath channel, the direct (or Line of Sight, LOS) signal arriving at the receiver can be significantly attenuated or blocked, and, in some cases, other multipath signals (echo signals) may reach the receiver with a higher power level than the attenuated LOS signal [8]–[10]. In such scenarios, the receiver may regard echo signal as the primary received signal, leading to ranging errors that vary from several to hundreds of meters [11]–[13]. Presently, environment-dependent multipath error has emerged as the dominant factor for positioning inaccuracies in receivers, and a definitive solution for mitigating multipath errors remains elusive [14].

Conventional GNSS receivers measure the time delay of the received signal by detecting the maximum peak in the sampled autocorrelation function (ACF) output envelope (ACF output) between the received signal and the replica signal generated by the receiver. In a multipath channel environment, the ACF output of the received BPSK-modulated signal has a triangular function-shaped peak in the time delay indicating each path. In a less multipath-affected environment close to the open sky, the first arrival path (FP) is likely to have the highest power, so the multipath error may be small at the receiver [15]. On the other hand, due to the BOC sub-carrier, the ACF output of a BOC-modulated signal that possesses side-peaks (i.e., fluctuating) can have a maximum peak in the ACF output that has a larger amplitude than the FP's main peak in multipath channel environments. As a result, the maximum peak may point to the echo path or side peaks. This situation frequently occurs in GNSS that use BOC modulation and other modulations derived from BOC (e.g., Composite BOC (CBOC), Time-division BOC (TMBOC), Alternative BOC (AltBOC)). Therefore, accurately detecting the FP time delay of the received signal in multipath channel environments and reducing the ambiguity of the autocorrelation of BOC modulation is a significant issue [16], [17].

Numerous studies have been conducted to reduce the ambiguity of the ACF main peak of BOC-modulated signals to enhance ranging accuracy [18]–[23]. The Bump-jumping tech-

*Corresponding author (e-mail: skong@kaist.ac.kr)

Sangjae Cho, Dong-Hee Paek, Hong-Woo Seok, and Seung-Hyun Kong are with the CCS Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Daejeon, Korea, 340-51 (e-mail: {sanjae, donghee.paek, hwseok}@kaist.ac.kr)

nique proposed in [18] detects the main peak by comparing the current peak to its neighboring peaks. However, this technique tends to be ineffective with low signal power. In [19], a BPSK-like approach was presented that addresses the BOC modulation as a sum of two BPSK modulations, resolving the ambiguity problem. Still, the BPSK-like approach compromises the signal tracking capability due to the narrowed width of the ACF main peak. [20] presented a technique subtracting the cross-correlation function of the BOC-modulated signal and the PN sequence from the ACF output, but this technique is only applicable to sine-phased BOC (BOCs) modulation. On the other hand, the techniques introduced in [21] and [22] are applicable to both BOCs and cosine-phased BOC (BOCc), but [21] doesn't effectively remove the side-peaks, and [22] shows worse signal tracking performance than the conventional ACF output, especially when the sub-carrier frequency is much greater than the chip rate. [23] presented a side-peaks cancellation (SC) approach that is simple and outperforms other SC techniques. However, it's susceptible to increased noise influence due to being generated by multiple subcorrelation's non-coherent integration.

Recently, [24] introduced DEEP Picker that utilizes Deep Learning (DL) to estimate parameters of each peak from inputs with multiple overlapping envelopes and to separate these peaks. It has been demonstrated that a data-driven approach can effectively analyze chaotically overlapped envelopes, and the consistency of its performance was also validated. The SC technique in this paper modifies the structure of DEEP Picker adapting it to be optimized for the SC task. This modification allows the technique to execute SC on the highly variable multipath channel environment's BOC ACF output without the necessity for multiple subcorrelation processes. With reference to a noise-free ACF output, there is no increase of noise. A comprehensive discussion on the SC used in BOC-FPDNet is provided in Section III.

To mitigate multipath errors, mathematical (deterministic) models based on [25]–[27] have been proposed. However, they still showed significant performance degradation in terms of positioning in NLOS channel environments. Subsequently, estimation techniques inspired by probabilistic models, such as the Space-alternation generalized expectation-maximization algorithm (SAGE) based on Maximum Likelihood (ML) and the Least-squares-based Iterative Multipath Super-resolution technique (LIMS) based on Super Resolution (SR), were introduced. Despite their advancements, these methods had the drawbacks of requiring immense computational costs due to the iterations and the need to pre-assume the number of multipaths. They also continued to show limitations in accuracy [28], [29].

With the advancement of computing technology, the use of Neural Networks (NN) has become prevalent in various fields [30]. NNs have demonstrated high performance in tasks such as classification [31]–[35], detection, and localization. In [15], the FPDNet, which mitigates multipath errors in GPS receivers, was introduced using MLP-Mixer, a highly efficient NN offering high performance relative to computational demands. However, FPDNet targets only BPSK-modulated signals that lack high amplitude side-peaks like BOC. This

limitation makes its application in GNSS restricted. Furthermore, since the final output of the NN is a time delay, there is insufficient conclusive evidence for detecting peaks in the ACF output that indicate the FP. As a result, FPDNet cannot handle to BOC-modulated signals, and there are concerns about the reliability of its outputs.

In this paper, we propose a first arrival path detection (FPD) technique, termed BOC-FPDNet, designed specifically for BOC-modulated signals characterized by the ambiguity of main peak. This is especially significant in multipath channel environments where the LOS path is frequently attenuated or obstructed, and multiple multipaths are received at the receiver. The proposed BOC-FPDNet adopts a two-stage architecture, sequentially performing SC and multipath time delay estimation (MDE) tasks. The MDE of BOC-FPDNet operates by analyzing the ACF output, detecting peaks indicative of paths, and subsequently estimating the time delay for each path. This MDE employs an artificial intelligence network that learns based on the correlations among the embedding vectors representing the input sequences, leveraging the Attention Mechanism (AM) [36]. The AM in MDE is utilized to learn the correlations between actual multipath time delays and the envelopes of the correlation outputs located at each path in the ACF output from multipath channel environments. As a result, in contrast to the previously introduced FPDNet [15], BOC-FPDNet does not merely regress to the FP's peak delay but performs time delay estimation for all paths within the ACF output. The advantages of our proposed BOC-FPDNet can be summarized as follows.

- BOC-FPDNet's SC operates using a data-driven algorithm, enabling side-peaks cancellation across all BOC modulation techniques.
- Unlike conventional SC, BOC-FPDNet's SC does not necessitate the computationally demanding subcorrelation, thereby preventing noise amplification.
- It deviates from the typical 2-stage structure, the MDE of BOC-FPDNet not only receives the refined input data, which is the output of SC, but also simultaneously takes in the original data (i.e., ACF output of the Original BOC-modulated signal), minimizing filtering loss due to SC.
- BOC-FPDNet's MDE, similar to the DL-based techniques in [15], does not solely detect the FP's time delay. Instead, it identifies the time delay of all received paths, enhancing the reliability of the output.

The structure of this paper is as follows. In Section 2, we introduce the mathematical representation of the ACF output, which is formulated by the combination of the BOC-modulated signal and subcorrelation. Additionally, the ACF output influenced by the sampling rate in multipath channel conditions is defined. Section 3 presents the two-tiered structure of BOC-FPDNet, constructed around Deep learning-based SC (DeepSC) and MDE, and further explains the technique used for detecting the earliest arriving path based on DeepSC outputs. In Section 4, we provide a comparative analysis of BOC-FPDNet's DeepSC with conventional SCs in terms of noise variations and overall accuracy. We also compare its efficacy with various neural networks constructed with

identical hyperparameters and both conventional mathematical and empirical model-based algorithms. Furthermore, we apply the BOC-FPDNet to actual GNSS receivers to verify positioning performance enhancements in multipath channel environments. Finally, Section 5 presents the final conclusions of this paper.

II. SYSTEM MODEL

This section introduces the mathematical expressions of the ACF output of the BOC-modulated signal, which is the input to BOC-FPDNet, and shows the side-peaks present in the ACF output. Furthermore, we derive the sampled ACF output envelope in terms of path delay, angular phase, and amplitude of the multipath.

A. Baseband BOC-modulated signal and ACF output

The baseband signal $s_{BB}(t)$ of the BOC modulation generated by the GNSS transmitter is defined as follows.

$$s_{BB}(t) = \sqrt{P} \sum_{i=-\infty}^{\infty} c_i p_{T_c}(t - iT_c) D_{\lfloor iT_c/T \rfloor}(t) s_i(t), \quad (1)$$

$$p_{T_a}(t) = \begin{cases} 1, & 0 \leq t < T_a \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where P represents the power of the signal, c_i is the i -th chip of the PN code with a period T , and is multiplied by $p_{T_c}(t)$ which has a rate $f_c (= n \times 1.023\text{MHz})$ and a width T_c . $D_j(t)$ represents the j -th bit of navigation data, and $\lfloor a \rfloor$ is the greatest integer not exceeding a . Within the PN code interval, since $D_j(t)$ is constant, it is approximated as 1 in this paper for the sake of simplicity. $s_i(t)$ is the BOC subcarrier and is defined as follows.

$$\begin{aligned} s_i(t) &= \text{sign}[\sin(2\pi f_{sc}(t - iT_c) + \theta_{BOC})] \\ &= \sum_{l=0}^{N-1} h_l p_{T_p}(t - iT_c - lT_p), \end{aligned} \quad (3)$$

where f_{sc} ($m \times 1.023\text{MHz}$) is subcarrier frequency. $s_i(t)$ can be either be sine-phased ($\theta_{BOC} = 0$) BOC ($\text{BOC}_s(m, n)$) or cosine-phased ($\theta_{BOC} = \pi/2$) BOC ($\text{BOC}_c(m, n)$) subcarrier depending on θ_{BOC} . Another representation of the BOC subcarrier, which is more intuitive and facilitates a concise derivation when expressing the autocorrelation function (ACF) of the BOC, is given by (3). Here N denotes the number of square waves within the subcarrier (e.g., N is 2 for $\text{BOC}_s(1, 1)$), $h_l \in \{-1, 1\}$ is the sign of the l -th pulse, and $T_p = T_c/N$ is the duration of the unit pulse of $s_{BB}(t)$. For $\text{BOC}_s(m, n)$, $\text{BOC}_c(m, n)$, the set (N, h_l, T_p) is specified as $(2m/n, (-1)^{(2m/n)i+l}, \frac{1}{2m \times 1.023\text{MHz}})$ and $(4m/n, (-1)^{(2m/n)i+\lceil l/2 \rceil}, \frac{1}{4m \times 1.023\text{MHz}})$, respectively, where $\lceil a \rceil$ is the smallest integer not smaller than a . To detect the presence or absence of signals emanating from GNSS satellites and to observe the time delay, the receiver correlates the received baseband signal $s_{BB}(t)$ with the replica code signal $s_{BB}(t + \tau)$ over the interval T . The normalized expected

output of the autocorrelation function can be represented as follows.

$$\begin{aligned} R_0(\tau) &= \frac{1}{PT} \int_0^T s_{BB}(t) s_{BB}^*(t + \tau) dt \\ &= \sum_{l=0}^{N-1} \sum_{j=0}^{\frac{T}{T_c}-1} \frac{1}{PT} \int_{(jN+l)T_p}^{(jN+l+1)T_p} s_{BB}(t) s_{BB}^*(t + \tau) dt \\ &= \sum_{l=0}^{N-1} \left\{ \frac{1}{N} \sum_{r=0}^{N-1} h_l h_r \Lambda_{T_p}(\tau + (l-r)T_p) \right\} \\ &= \sum_{l=0}^{N-1} S_l(\tau), \end{aligned} \quad (4)$$

where

$$\Lambda_a(\tau) = \begin{cases} 1 - \frac{|\tau|}{a}, & |\tau| \leq a \\ 0, & |\tau| > a \end{cases} \quad (5)$$

is a triangular function, and $S_l(\tau)$ is called the l -th sub-correlation function. Note that the normalized ACF output includes N sawtooth-shaped functions, $\{S_l(\tau)\}_{l=0}^{N-1}$, each with different phases. Due to these characteristics, the ACF output has several side-peaks, leading to increased ambiguity of the main peak compared to the ACF output of a BPSK-modulated signal. Additionally, as the subcarrier frequency rises (i.e., T_p decreases and N increases), while the main peak of the BOC-modulated signal's ACF output becomes sharper. Concurrently, the count of $S_l(\tau)$ grows, leading to a proportional increase in the number of side-peaks. The BOC-modulated signal produces the ACF output with several side-peaks even in open sky conditions, resembling a signal received in a multipath channel environment. Consequently, in a multipath channel scenario, detecting the main peak is substantially more challenging than with a BPSK-modulated signal. Hence, to minimize the multipath error of the BOC-modulated signal, it is imperative to first alleviate the ambiguity associated with the ACF output's side-peaks.

B. ACF output of multipath environment

In the multipath channel environment, when a wireless signal reaches the receiver, its time domain Channel Impulse Response (CIR) $h(t)$ can be modeled as follows.

$$h(t) = \sum_{k=0}^{K-1} C_k \delta(t - \tau_k), \quad 0 \leq t < T_{co}, \quad (6)$$

where K is the number corresponding to the total number of paths,

$$C_k = a_k e^{j\theta_k}, \quad (7)$$

represents the complex channel coefficient of the $(k+1)$ th path, where $a_k (\leq 1)$ is the amplitude, and θ_k is the angular phase. $\tau_k (> \tau_{k-1})$ is the dirac delta function, and $\tau_k (> \tau_{k-1})$ denotes the delay of the $(k+1)$ th path. Using (6), the baseband received signal in the continuous-time domain can be found as follows

$$y_{BB}(t) = s_{BB}(t) * h(t) + n(t), \quad 0 \leq t < T_{co}$$

$$= \int_{-\infty}^{\infty} s(\tau)h(t+\tau)d\tau + n(t), \quad (8)$$

where ‘*’ denotes the convolution operation, and $n(t)$ represents the complex additive white Gaussian noise (AWGN) process with a two-sided power spectral density (PSD) of $N_0/2$. The baseband received signal $y_{BB}(t)$ is despread through a correlation process with the BOC modulated replica signal $s_{BB-R}(t)$ generated at the receiver. In this context, the normalized ACF output in the continuous-time domain can be expressed as follows

$$R(\tau) = \frac{1}{T_{co}} \int_0^{T_{co}} y_{BB}(t)s_{BB-R}^*(t+\tau)dt$$

$$= \sum_{k=0}^{K-1} C_k R_0(\tau + \tau_p) + \omega(\tau), \quad (9)$$

where $\omega(\tau)$ denotes zero-mean gaussian process with autocorrelation $E\{\omega(\tau)\omega^*(\lambda)\} = (\frac{N_0}{2T_{co}})R_c(\tau - \lambda)$, R_0 is noise-free BOC modulated signal ACF. Sampled ACF output envelope $R(nT_s)$ with sampling period T_s and the sampling frequency $f_s(f_s = 1/T_s)$ becomes

$$R(nT_s) = \begin{cases} \sum_{k=0}^{K-1} C_k R_0(nT_s - \tau_k) + \omega(nT_s), & \left\lfloor \frac{\tau_0 - T_c}{T_s} \right\rfloor \leq n \leq \left\lfloor \frac{\tau_{K-1} + T_c}{T_s} \right\rfloor \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

where n is the sample index defined within the range where the ACF output $R(nT_s)$ is non-zero, and τ_k can be expressed as the quotient $d_k(d_k \geq 0)$ and the remainder $\epsilon_k(0 \leq \epsilon_k < T_s)$ when divided by T_s .

$$\tau_k = d_k T_s + \epsilon_k, \quad (11)$$

$$d_k = \left\lfloor \frac{\tau_k}{T_s} \right\rfloor, \quad (12)$$

where

$$\epsilon_k = \tau_k - \left\lfloor \frac{\tau_k}{T_s} \right\rfloor T_s. \quad (13)$$

In other words, d_k represents the path delay of the k -th path in sample units, while ϵ_k means the residual delay of the k -th path induced by sampling. In LOS channel, the magnitude $|C_k|$ of the first path ($k=0$) becomes the maximum value of $R(\tau)$ as (9). When $\epsilon_0 = 0$, the maximum point of $R(\tau)$ (9) aligns with that of $R(nT_s)$ (10). However, typically, $\epsilon_k \neq 0$, leading the maximum value of the ACF output $R(nT_s)$ to be less than the peak value of the continuous ACF output $R(\tau)$. Due to this, while in a multipath environment the arrival time of the FP in the ACF output may become more accurate as T_s decreases, a higher sampling rate f_s comes with the downside of increasing the complexity of the receiver.

III. BOC-FPDNET: 2-STAGE MULTIPATH TIME DELAY ESTIMATION

In this section, we introduce the BOC-FPDNet, which follows a two-stage structure where two modules, SC and MDE, operate sequentially. The two-stage architecture is commonly found in object detection algorithms [37], [38]. In the first stage of these object detection frameworks, a rough (coarse) location of the object is identified, while the second stage refines this position to obtain a more precise (fine) location along with additional information, such as classification. We tailored our approach into this two-stage design to cater to the BOC-modulated signals that require tasks like reducing the ambiguity of main peak in a multipath channel environment and detecting the time delay of the FP. Similar to conventional detectors, in the first stage (DeepSC), SC is executed to predict the main peak of the ACF, while in the second stage (MDE), the previously predicted main peak data is used to anticipate the delay of each received signal path and their relative amplitudes. We introduce the first stage of BOC-FPDNet, DeepSC, in Subsection III-A, and the second stage, MDE, in Subsection III-B.

A. DeepSC: Deep learning-based approach for side-peaks Cancellation

As mentioned in [18]–[23], the key to enhancing the positioning accuracy of BOC signals is minimizing the ambiguity of main peak in the correlation output. SC is the most representative technique for reducing this ambiguity, by merging multiple subcorrelation outputs of BOC signals to cancel out the side-peaks. For instance, the SC for BOC(n, n) (e.g., BOC(1,1) in this paper) proposed in [23] is executed as follows.

$$|R_{SC}| = |R_{sub1}| + |R_{sub2}| - |R_{sub1} + R_{sub2}|, \quad (14)$$

where R_{SC} is the ACF output of BOC(n, n) with SC applied, and R_{sub1} and R_{sub2} represent the outputs of the subcorrelation between BOC(1,1) with a time delay of $\pm 0.5T_c$ and BPSK(1). In the SC process, the influence of noise increases due to the non-coherent integration as in (14). Moreover, as the order of BOC increases, more subcorrelations are required, significantly intensifying the effects of noise [39]. Consequently, conventional SC techniques are vulnerable in noise-intensive channel environments, and applying them to BOC modulations with high subcarrier frequencies ($m \gg n$) is challenging in real multipath channel scenarios.

Inspired by [24], DeepSC, as depicted in Fig. 1, employs multiple convolution layers to perform side-peaks cancellation on the envelopes of each path of the ACF output over a sliding window. DeepSC adopts a multi-task structure similar to the approach presented in [24]. Unlike [24], we address the SC task and estimate the time delay $\hat{\tau}_k$ and amplitude $\hat{a}_k$ for each path. Here, the SC differentiates every sample point in the sampled BOC ACF envelope as $R_0(\tau)$ or $\omega(\tau)$, finally producing BOC ACF output where only the side-peaks are eliminated. Additionally, the tasks of estimating $\hat{\tau}_k$ and $\hat{a}_k$ serve as auxiliary tasks that aid the DeepSC's

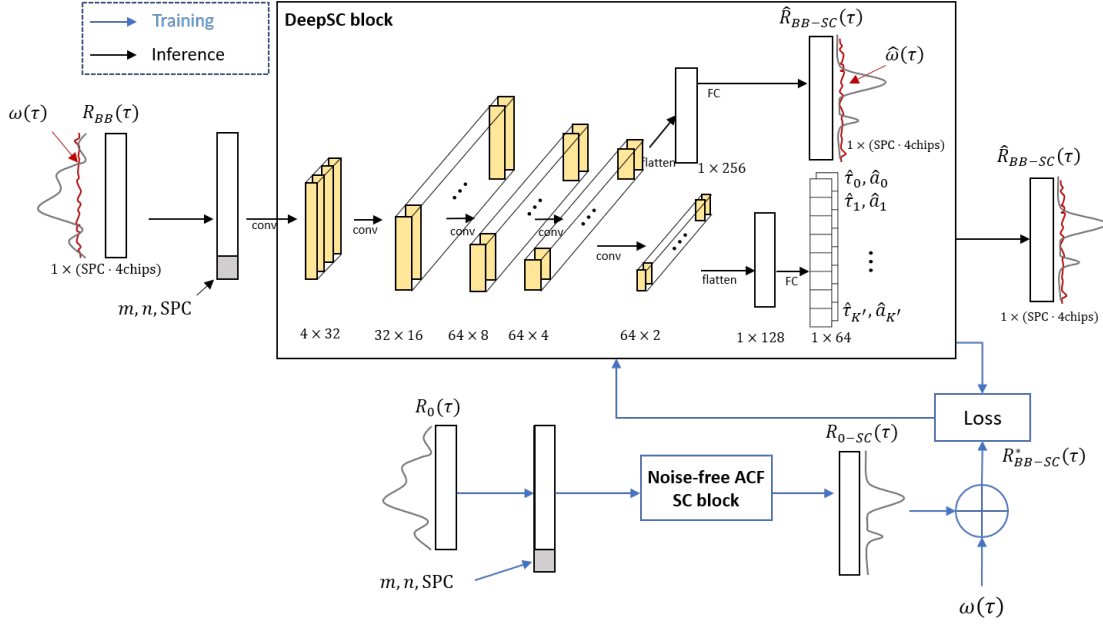
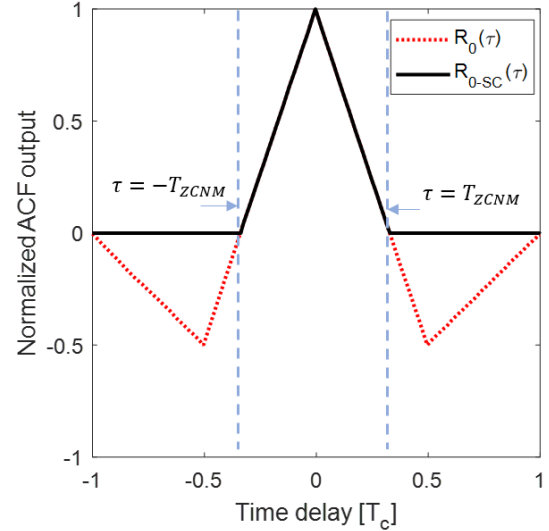


Fig. 1. Overall structure of DeepSC.

Backbone network in extracting unique features of the main peak, indicating each path in the ACF output.

The overall structure of DeepSC, as shown in Fig. 1, accepts an ACF output of length samples per chip (SPC) $\times 4T_c$ (i.e., $[-1.5T_c, 2.5T_c]$) considering the actual multipath delays and outputs a corresponding size $\hat{R}_{BB-SC}(\tau)$. DeepSC model consists of five hidden layers, one maximum pooling layer, and two parallel output layers. The total number of trainable parameters is 48k for 4SPC, 56k for 8SPC, and 73k for 16SPC. The size of the feature map for each layer is demonstrated in Fig. 1. Although not shown in Fig. 1, we employed Max pooling to effectively compress the features of the final convolution layer of DeepSC. To realize a non-linear SC, we incorporated a ReLU activation function in each layer, following the approach in [40]. As with [24], the loss functions for the regressor and side-peaks cancellation were implemented using Mean Squared Error (MSE) and Cross-entropy, respectively. The training of DeepSC utilized the Adam optimizer with a learning rate of 0.001, as described in [41]. We applied two manipulations to the conventional BOC ACF output in DeepSC to alleviate its ambiguity. Firstly, to label the output of DeepSC, the reference side-peaks cancelled ACF for DeepSC was generated. This was done by setting the correlation output from the zero-crossing near the ACF main peak (ZCNC) T_{ZCNC} up to one chip to zero. As a result, it produced a noise-free ACF output with cancelled side-peaks, denoted as $R_{0-SC}(\tau)$. The ideal output of DeepSC can be expressed as follows.

$$R_{0-SC}(\tau) \begin{cases} R_0(\tau), T_c < |\tau| \leq T_{ZCNC}, \\ 0, T_{ZCNC} < |\tau| \leq T_c, \end{cases} \quad (15)$$

Fig. 2. $R_0(\tau)$ and $R_{0-SC}(\tau)$ of BOC(1,1)-modulated signal.

where

$$T_{ZCNC} = \pm T_c \times (4m - n)^{-1}. \quad (16)$$

Fig. 2 presents examples of $R_0(\tau)$ and $R_{0-SC}(\tau)$ for BOC(1,1). For BOC(1,1), T_c is $0.98\mu\text{sec}$, and T_{ZCNC} is $0.33T_c$. Following (15), the $R_{0-SC}(\tau)$ for BOC(1,1) only allows the peak within the range of the main peak $T_c < |\tau| \leq 0.33T_c$ in the $\pm 1T_c$ interval, as shown in Fig. 2.

Secondly, by adding $\omega(\tau)$ to $R_{0-SC}(\tau)$, we generated the ACF output $\hat{R}_{BB-SC}(\tau)$ where only the side-peaks are cancelled from $R_{BB}(\tau)$ without altering the size of the noise, different from (14). Fig. 3 shows examples of the ACF output $R_{BB}(\tau)$ of the received baseband BOC(1,1) signals and the

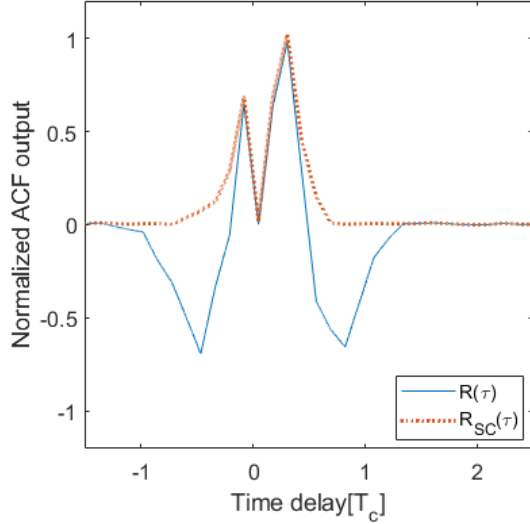


Fig. 3. $R_{BB}(\tau)$ and $R_{BB-SC}^*(\tau)$ of BOC(1,1)-modulated signal in multipath channel.

side-peak cancelled ACF output $R_{BB-SC}^*(\tau)$ over the chip delay τ in the interval $[-1.5T_c, 2.5T_c]$. $R_{BB-SC}^*(\tau)$ only includes the main peak and $\omega(\tau)$ from $R_{BB}(\tau)$ because only the side-peaks in the range $T_{ZCNM} \leq |\tau| \leq T_c$ are cancelled. Since extracting accurate noise and paths from real received signals is realistically challenging, we generate them using a virtual multipath channel environment. Thus, we can produce abundant data for the empirical model, DeepSC, across various multipath channel scenarios. A detailed description of the virtual multipath channel environment is covered in Section IV-A.

In [24], which shares a similar architecture with DeepSC, experimental results have confirmed the advantage of employing separate neural network components in the output layer to predict two distinct types of peak parameters: the estimated positions and amplitudes of peaks, and the classified positions in the input. This illustrates that by using different labels for these two peak types, the same objective can be achieved with enhanced performance [42]. It is important to note that we adopted the multi-task approach to enhance the feature extraction capability of DeepSC's backbone network (Convolutional Neural network layers). In practice, we only utilize the output that filters the Main peak (i.e., $\hat{R}_{BB-SC}(\tau)$) and feed it into the 2nd stage of the BOC-FPDNet, instead of using both types of outputs.

B. MHA-based Multipath time delay estimation

The 2nd stage of the BOC-FPDNet proposed in this paper is the Multipath time delay estimation (MDE) based on multi-head attention (MHA) [36] to detect the time delay of each path within the ACF output of the received signal in a multipath channel environment. The Convolutional Neural Network (CNN) [43], [44], which leverages convolution operations to extract the primary features of local samples within the ACF output, requires computations in deeper layers to encode global

samples. As a result, CNN experiences a resolution loss of local samples during the encoding process of global samples. Since the DeepSC, which identifies the global sample shape (i.e., sawtooth-shape function shape) to cancel the side-peaks, considers the global sample shape more crucial than the local sample shape (e.g., $\omega(\tau)$), employing a CNN is appropriate. However, for MDE, whose main objective is to judge and detect samples pointing to the local information, the main peak, in $\hat{R}_{BB-SC}(\tau)$, it is vital to concurrently grasp both the global sample shape and the shape of the local samples.

In the realm of natural language processing problems, where understanding the meaning of local words (or samples) within the global context of input data is crucial, Self-attention [45] is predominantly utilized. Self-attention is a neural network that computes correlations between a local word in a sequence and every other word, ensuring the extraction of only pertinent words without any loss of meaning. Multi-Head Attention (MHA) [36] dissects the self-attention mechanism into multiple heads, providing an advanced capability to analyze high-dimensional correlations and dramatically reducing computational overhead compared to [45]. In other words, MHA simultaneously considers various associations of input words, proving to be exceptionally useful in detecting and understanding intricate patterns, surpassing conventional networks like Multi-layer Perceptron (MLP), Residual Neural Network (RNN), and CNN. Therefore, we propose the use of MHA, apt for discerning the shape of global samples without compromising the resolution of local samples, in the 2nd stage of BOC-FPDNet.

Fig. 4 presents the overall structure of the BOC-FPDNet. In the 1st stage, the ACF output of the conventional BOC-modulated signal in the complex domain and the final output of DeepSC are non-coherently integrated and then concatenated. To merge the information from both data sources, they are passed through a Fully-connected (FC) layer and transformed into a feature map of size $SPC \times 4 \times N_{chips}$. To enhance the resolution of the input data, this feature map is expanded into a vector of size N_{ts} (set to 16,000 in this paper). Thus, as illustrated in Fig. 4, both the original BOC-modulated signal and each ACF output with SC applied are simultaneously fed into the MDE constructed using the MHA mechanism, minimizing the filtering loss due to DeepSC. For the MDE, considering its application in low-cost receivers, the embedding dimension, number of heads, and expansion factor were chosen to be 16, 4, and 8, respectively. The input vector to the MDE passes through layer normalization, MHA, and a Feed Forward layer (comprising a Linear layer, ReLU activation, and another Linear layer). Ultimately, the vector is translated into Positive (indicating the presence of the path) and Negative (indicating the absence of the path) (P/N) samples of size 2×640 . The argmax function then represents the sample indicating the time delay of each path. It should be noted that in order to detect sparse time delay samples throughout the entire sample set, we aimed for a more reliable output than a single vector by performing a different task called P/N sample detection. For the detection of P/N samples in the MDE, we defined the loss function as follows.

$$L_{total} = \alpha L_{pos} + \beta L_{neg}, \quad (17)$$

$$L_{pos} = -\frac{1}{N_{pos}} \sum_i^{N_{pos}} y_{pos,i} \log(\hat{y}_{pos}), \quad (18)$$

$$L_{neg} = -\frac{1}{N_{neg}} \sum_i^{N_{neg}} y_{neg,i} \log(\hat{y}_{neg}), \quad (19)$$

where y_{pos} and y_{neg} are corresponding P/N samples. (17), (18), and (19) are based on the Class-Balanced Loss Based on the Effective Number of Samples (Class-balanced loss) [46]. [46] proposed a method to address the issues with class-unbalanced data. Imbalanced data refers to cases where the number of samples for each class within the dataset differs significantly. In this paper, since only a tiny fraction of the entire ACF output samples indicate paths, samples pointing to the path are considered as y_{pos} as per (18), while other samples are considered as y_{neg} according to (19). The weight for each category is set as the inverse of the sample count, leading to higher weights for classes with fewer samples and lower weights for classes with more samples. Ultimately, adjusting these weights has the effect of multiplying the loss of each class when computing the loss function, enabling a more focused attention on minority class samples (i.e., samples indicating paths). Finally, the MDE of BOC-FPDNet returns samples that match the time delay of each path within the ACF output, and we extract only the foremost index from its output. This index is converted from the sample level to the time level to determine the time delay. When the foremost index is i_{FP} , the time delay corresponding to that index is given by $i_{FP} \times (4T_c/N_{sample})$, where c and N_{sample} represent the speed of light and the total length of the output samples, respectively. For instance, a conventional GPS receiver with the sampling rate $f_s = 4.092\text{MHz}$ has an estimation error of approximately 36.7m due to sampling resolution. However, since the final output size of BOC-FPDNet is N_{patch} , the error is reduced to 0.93m.

The network training is conducted on an NVIDIA RTX 3080, using the AdamW optimizer [41] with a learning rate of 0.001. Xavier normal is employed as the kernel initializer [47].

IV. PERFORMANCE EVALUATION

In this section, we present two principal performance evaluations. Firstly, we evaluate the side-peaks cancellation efficacy of the first stage, DeepSC. Subsequently, leveraging simulations modeled on various multipath channel conditions and actual field tests, we assess the time delay estimation capability for the initial arriving path in the BOC-FPDNet's second stage, MDE. While this section predominantly addresses the BOC(1,1) modulation that is representative of BOC modulations.

For a more authentic performance assessment, we incorporate the distortion effects on the ACF output induced by

the pre-correlation bandwidth (PCBW), as described in [48]. Such distortions typically manifest as rounded peaks in the ACF output. It is well-understood that signals traversing a bandlimited filter experience a loss in their high-frequency components, leading to a smoothening of sharp ACF peaks. Consequently, accounting for the effects of PCBW becomes imperative when one is engaged in estimating multipath time delays via ACF. Given that the distortions from PCBW are orthogonal to noise and multipath effect, our model sidesteps noise considerations and presumes $y_{BB}(t)$ to signify a single path scenario, that is, $K = 1$.

Given that $y_{BB}(t)$ and the impulse response of the ideal low-pass filter $h_{LPF}(t)$ with PCBW B are Fourier transformed into $Y_{BB}(f)$ and $H_{LPF}(f)$, respectively, the Fourier transform of the ACF output $R(\tau)$ can be expressed as follows.

$$\begin{aligned} R_{BB}(f) &= Y_{BB}(f)S^*(f) \\ &= H_{LPF}(f)|S(f)|^2, \end{aligned} \quad (20)$$

where

$$H_{LPF}(f) = \begin{cases} 1, & |f| < B, \\ 0, & \text{otherwise,} \end{cases} \quad (21)$$

is a rectangular function with a bandwidth B , and $S(f)$ represents the Fourier transform of the BOC-modulated signal produced by the receiver. The bandlimited ACF output $R_{BL}(t)$ can be approximated as follows.

$$\begin{aligned} R_{BL}(t) &= \mathcal{F}^{-1}\{H(f)|S(f)|^2\} \\ &\cong B \times \text{sinc}(\pi Bt) * R_0(t). \end{aligned} \quad (22)$$

In the above, $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse Fourier transform, $R_0(t)$ represents the ideal ACF output in the infinite bandwidth, and $\text{sinc}(\cdot)$ is defined as $\text{sinc}(x) = \sin(x)/x$. Based on the analysis in [48], to approximate $R(t)$ using a functional representation, one can consider the BOC-modulated ACF output having a triangular function shape, i.e., $R_0(t) = \Lambda_{T_c}(t)$, then $R(t)$ is given by

$$R(t) \cong \beta \int_{-1}^1 \text{sinc}(\beta\tau) R_0(t - \tau) d\tau, \quad (23)$$

where $\beta = BT_c$ so that $\beta = 0.5$ is for the receiver that has B equal to the chip rate of the spreading code. Fig. 5 shows the ACF outputs of the bandlimited BOC(1,1) signals for $\beta = 0.5, 1, \text{and } \infty$.

Furthermore, when β is less than 1, it was observed that the amplitude of the main peak decreases, and its shape becomes less pronounced or less sharp. In the simulation, we employed three typical sampling frequencies for BOC(1,1) — $4/T_c$, $8/T_c$, and $16/T_c$ (equivalent to 4, 8, and 16 samples per chip (SPC) and $\beta = 4, 8$, and 16). Generally speaking, the code phase search resolution of the acquisition function for the BOC-modulated signal is $0.5T_{ZC/NM}$. While actual receivers usually exhibit the FP delay error less than $0.5T_{ZC/NM}$ in an open sky environment, we assumed the FP's time delay τ_0 to be uniformly distributed over the interval $[-T_c, T_c]$ to

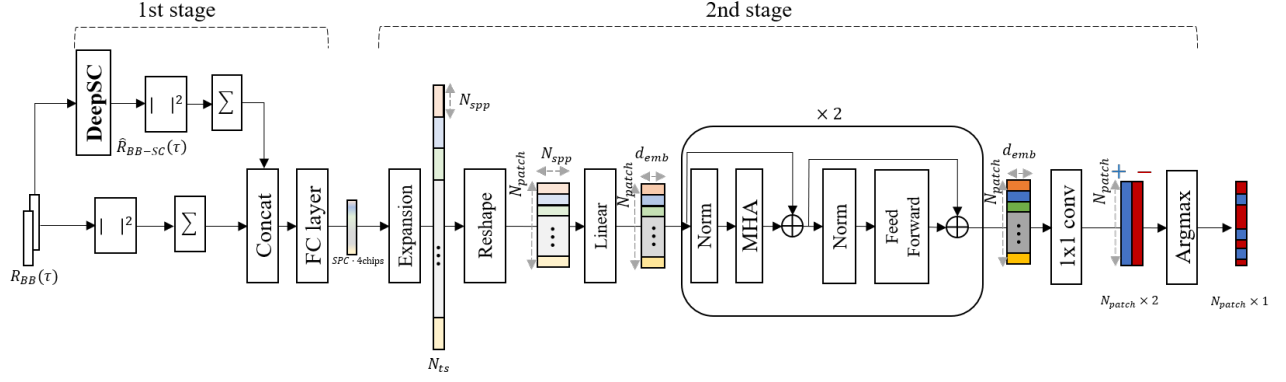


Fig. 4. Overall structure of BOC-FPDNet.

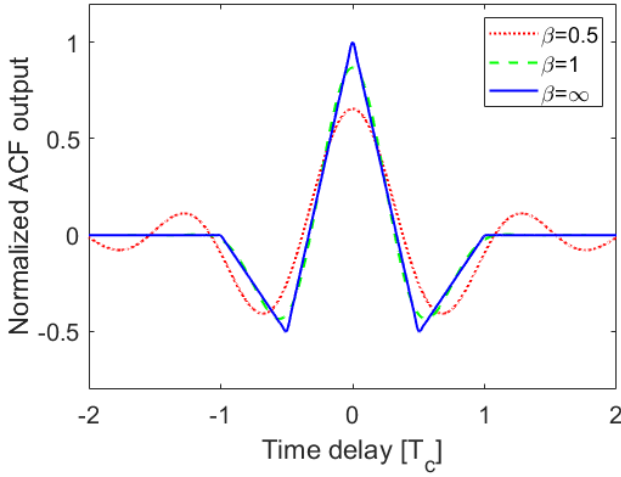


Fig. 5. ACF output for bandlimited BOC(1,1)-modulated signals.

cater to all GNSS signals, including BPSK. As observed in [29], [49], multipath delays are rarely greater than T_c . Thus, we can indirectly infer that the actual FP's time delay τ_0 is not earlier than $-0.5T_c$ from the detected FP delay, and the last path delay $\tau_{(K-1)}$ is not later than $1.5T_c$ from the detected FP delay. Considering that the ideal path's ACF output encompasses a range from $-T_c$ to $+T_c$ around the peak and that $\tau_{(K-1)} - \tau_0 < 2T_c$, the ACF output influenced by the incoming signal paths resides within a window width of $4 \text{ chips} \times T_c$. The total number of samples N_{sample} within the window equates to $4\text{chips} \times (T_c/T_s)$.

For a fair performance assessment, the assumptions integrated into our simulations align with those presented in [15] and summarized as follows. 1. DNN-driven techniques assume $T = 1\text{ms}$. However, for LIMS and SAGE, as indicated in [29], the assumption stands at $T = 10\text{ms}$. It means that DL-based techniques take a more challenging scenario than SR-based techniques. 2. Considering hardware constraints that lock the sampling rate as a fixed parameter, distinct DNN-centric methods are trained for different SPCs. Specifically, for BOC(1,1)-modulated signals, separate DNN-driven models are trained and subsequently tested for each of the $\text{SPC} = (4, 8, \text{and } 16)$.

3. As T augments, resulting in a tenfold rise in the SNR, LIMS and SAGE inherently possess an advantage concerning noise. Furthermore, the DNN-driven approaches are trained collectively on all virtual multipath channel environments discussed in this paper and are tested against specific channels. Consequently, the training and test datasets are distinctly set apart. The training dataset comprises 4 Carrier to noise ratios (C/N_0), 5 multipath channel environments, and a total of 10,000 Monte Carlo realizations (each with a distinct set of 100 sampling offsets and time delays), aggregating to 600,000 instances in total. The test dataset, on the other hand, is 10% of the training set, tallying up to 60,000 instances."

A. Side-peaks cancellation performance

As demonstrated in [18]–[22], the performance evaluation of conventional SC techniques primarily revolved around applying them to Acquisition or Tracking and then comparing the ranging error of the receivers. Since this doesn't directly evaluate the reduction of ambiguity in the main peak of the actual received signal's ACF output, we assess the SC performance of DeepSC based on the similarity between the reference side-peak cancelled ACF output, $R_{BB-SC}^*(\tau)$, for a given C/N_0 , and the output of DeepSC, $\hat{R}_{BB-SC}$.

Table I presents the SC performance results trained on a total of 200,000 ACFs of BOC(1,1)-modulated signals (for BOCs and BOCc with different SPCs). Note that DeepSC requires no additional adjustments regardless of BOCs and BOCc. The similarity and noise variation between $R_{BB-SC}^*(\tau)$ and the output from DeepSC, $\hat{R}_{BB-SC}(\tau)$, were evaluated across all SPCs and C/N_0 s. Table I shows negligible performance differences between these two modulations. On average, DeepSC has 93.80% similarity and 0.74dB noise variation. As the similarity encompasses results over all noise-inclusive regions, considering the amount of noise variation, one can infer that DeepSC sufficiently predicts the main peak after canceling the side-peaks. It is evident that as SPC increases, DeepSC benefits from more samples, enhancing the similarity performance. However, with an increase in SPC, more noise is introduced, leading to a slight increase in noise variation. Similarly, as C/N_0 decreases, the ambiguity in the ACF output's main peak increases due to the elevated impact of noise, escalating the noise variation. Nonetheless,

TABLE I
SIDE-PEAKS CANCELLATION PERFORMANCE OF DEEPSC FOR BOC(1,1)

SPC	C/N ₀ (dB-Hz)	35		40		45		50	
	Modulation	Similarity (%)	NV (dB)	Similarity (%)	NV (dB)	Similarity (%)	NV (dB)	Similarity (%)	NV (dB)
4	BOCs(1,1)	89.21	1.14	91.16	0.92	91.65	0.46	94.03	-0.43
	BOCc(1,1)	89.16	1.20	92.00	0.94	92.13	0.49	94.62	0.22
8	BOCs(1,1)	91.39	1.34	94.71	0.77	95.98	0.51	97.12	0.14
	BOCc(1,1)	90.11	1.36	94.15	0.81	95.62	0.51	96.80	0.20
16	BOCs(1,1)	92.22	1.41	95.54	0.93	97.87	0.86	96.98	0.35
	BOCc(1,1)	91.41	1.45	94.71	0.92	96.10	0.93	96.59	0.41

TABLE II
MULTIPATH CHANNEL PARAMETERS

Path	Channel-2			Channel-3			Channel-4		
	$k = 0$	$k = 1$	$k = 0$	$k = 1$	$k = 2$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
Relative Power (dB)	-3	-3	0	-5	-10	-7	-7	0	-2.2
Relative Delay	0	$T_c/2$	0	$[0.1T_c, 0.3T_c]$	$[0.3T_c, 0.5T_c]$	0	$[0.1T_c, 0.3T_c]$	$[0.3T_c, 0.4T_c]$	$[0.5T_c, 0.7T_c]$
Random Phase	Y								

this variation is minute compared to the main peak, implying that predicting the main peak isn't significantly problematic. Though not delved into in this paper, DeepSC is capable of performing an effective SC task for the single ACF output in the range $[-1.5T_c, 2.5T_c]$ through simple matrix multiplication, making its computational requirements considerably less compared to conventional SC techniques that execute multiple subcorrelations over the entire PN code range. Note that unlike DeepSC, conventional SC techniques require an increased number of subcorrelations as the BOC's subcarrier increases. The evaluation of DeepSC performance for higher subcarrier frequency (i.e., BOC(kn, n)) is discussed in detail in Appendix. Consequently, in environments such as urban settings, where the impact of noise is significant and multipath channels are prevalent, DeepSC emerges as the most effective and suitable SC technique for BOC-modulated signals.

B. First arrival path delay estimation

The MDE of BOC-FPDNet estimates the time delay of each multipath component received. However, to minimize the receiver's ranging error, as previously discussed, it is crucial to accurately estimate the time delay of the first arrival path. Therefore, the performance evaluation for BOC-FPDNet assumes the estimated FPD $\hat{\tau}_0$ as the time delay of the foremost path among all the multipath time delays, and it is compared with the reference FPD τ_0^* . Table II summarizes the three types of virtual multipath channel environments on which testing was performed for performance comparison. These channel environments are consistent with those utilized for performance evaluation in [15]. The term Channel- K signifies a channel environment comprising K paths. Both Channel-2 and Channel-4 represent NLOS channel environments where the FP is affected by fading. Relative delay is referenced to the time delay of FP, τ_0^* . Though not specified in Table II, in all simulations, C/N_0 comprises four values ranging from 35 to 50dB-Hz in increments of 5dB-Hz. As illustrated in Table I, the tendencies for BOCs(1,1) and BOCc(1,1) are remarkably congruent. Accordingly, we show the mean FP delay estimation performance for both BOCs(1,1) and BOCc(1,1) modulations.

TABLE III
COMPARISON OF MODEL SIZE OF DL-BASED TECHNIQUES

	Number of Parameters
VGGNet [15]	2.0M
ResNet [15]	2.1M
FPDNet (MLP-Mixer) [15]	1.14M
BOC-FPDNet (DeepSC+MDE, 4SPC)	53.6k (48k+5.6k)
BOC-FPDNet (DeepSC+MDE, 8SPC)	61.7k (56.1k+5.6k)
BOC-FPDNet (DeepSC+MDE, 16SPC)	78.3k (72.7k+5.6k)

In general, for signals captured in open areas (i.e., LOS channel conditions), the ACF output, despite the presence of side-peaks, shows a distinct main peak. This makes it feasible to adequately detect the signal using a multiple correlator-based discriminator. However, in multipath channel environments, where the relative delay is all within $1T_c$ as shown in Table II, there's a significant degradation in performance. Fig. 6 shows the FP delay estimation performances for Channel-2, Channel-3, and Channel-4 using SR-based techniques presented in [28], [29], and DL-based techniques presented in [15], along with BOC-FPDNet. SR-based techniques are assumed for 10SPC, and DL-based techniques are under assumptions of 4SPC, 8SPC, and 16SPC. Table III summarizes the number of learnable parameters for DL-based techniques. SR-based techniques perform FPD with significantly fewer parameters compared to DL-based techniques, but it's difficult to compare their exact complexity since the iteration in SR-based techniques is determined by convergence, unlike DL-based techniques that perform a single iteration. BOC-FPDNet has the fewest number of parameters among other DL-based techniques, with up to approximately 37 times fewer.

For LIMS and SAGE, designed based on mathematical rules to detect peaks with infinite precision, they demonstrate a consistent RMS error of over 80m, failing to accurately detect the First peak despite changes in C/N_0 , unlike the BPSK-modulated signal scenario shown in [15]. This error essentially emerges when targeting the midpoint between the

main peak and the second peak, roughly at $0.25T_c$. This outcome indicates that, due to the influence of side-peaks in the ACF output of the BOC-modulated signal, there is a significant increase in ambiguity related to the main peak of the ACF output. In multipath channel conditions, conventional mathematically modeled techniques can introduce positioning errors of several hundreds of meters when receiving a BOC-modulated signal.

However, DL-based approaches like VGGNet, ResNet, and FPDNet show a reduction in RMS error with increasing noise and an increase in SPC. This is due to their ability to recognize diverse features detected in the ACF output, grounded in data-driven insights (i.e., extensive experience across a broad range of ACF outputs). Nonetheless, compared to the results shown in [15] for the BPSK-modulated signal, VGGNet, ResNet, and FPDNet all exhibit relatively higher errors due to the influence of side-peaks. FPDNet, in particular, demonstrates lower error compared to CNN-based methods, given its ability to extract global features from the ACF output. This suggests that empirical learning of the ACF output, including side-peaks, leads to improved performance. Among CNN-based methods, ResNet outperforms VGGNet slightly, attributable to ResNet's use of residual blocks which help preserve the resolution of the ACF output even in deeper layers. In comparison with other DL-based techniques, the proposed BOC-FPDNet consistently exhibits lower error across all SPCs. This is because BOC-FPDNet, unlike other DL techniques, reduces ambiguity related to side-peaks in the ACF output through DeepSC and bases its findings on precise localization of MDE. However, VGGNet, ResNet, and FPDNet demonstrate better performance than BOC-FPDNet only when C/N_0 is 50dB-Hz. The reason is that, excluding BOC-FPDNet, other DL methods regress the final output to the FP's time delay, allowing for estimations with unrestricted output resolution, thus enabling extremely accurate predictions. Through the analysis in Fig. 6, the significant advantages of BOC-FPDNet become evident. Unlike Channel-2 and 3, even in Channel-4, where the power of the First arrival path is significantly attenuated, DL-based techniques, including BOC-FPDNet, showcase superior FPD performance compared to formula-based techniques. Essentially, data-driven approaches are determined by the diversity and scale of the data. Thus, given ample and diverse data, we can factually assert that DL-based techniques, for all channel environments encompassed by the dataset's distribution, offer superior performance characterized by greater generality and higher accuracy compared to formula-based methods. To compare the performance of BOC-FPDNet in a stochastic multipath channel simulation environment reminiscent of urban areas, we employ the Land Mobile Multipath Channel (LMMC) model standardized by the International Telecommunication Union (ITU) as ITU-R P.681.7 [50]. This channel model generates the Channel Impulse Response (CIR) in urban areas based on Ray tracing, allowing the user to manually select specific urban settings (e.g., trees, utility poles, building spacing) and consequently produce the CIR according to a probability distribution that matches the specified environment. Consistent with the consideration in [15], we have chosen two identical virtual city settings, namely Urban and Dense urban.

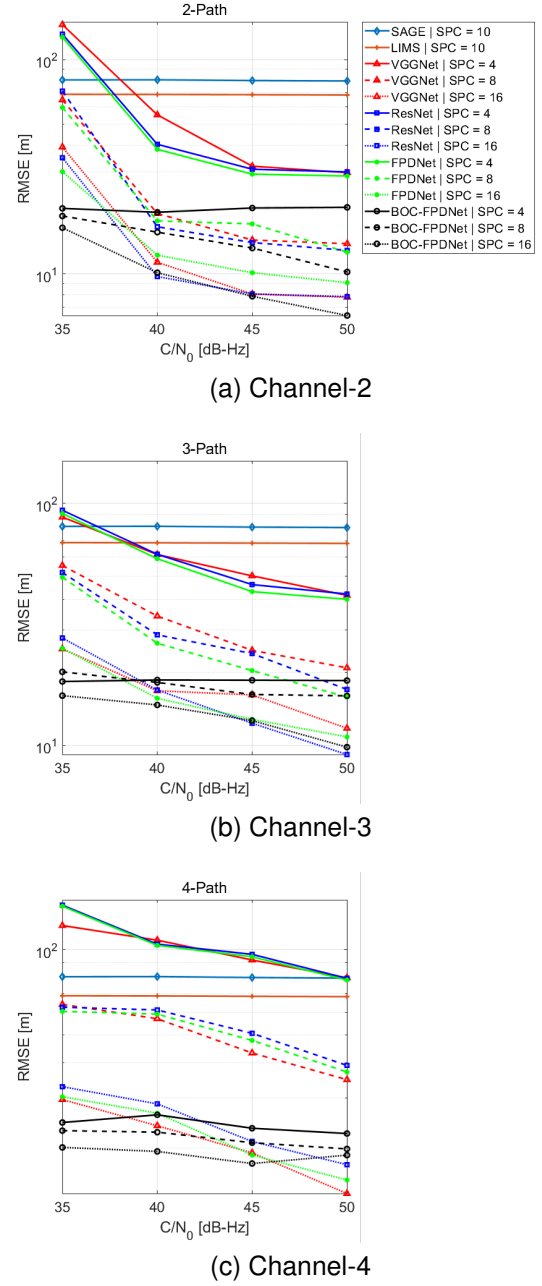


Fig. 6. RMS error of Channel- K ($K = 2, 3, 4$).

The parameters for these two environments are summarized in Table IV.

Fig. 7 shows the FPD performance results of VGG, ResNet, FPDNet, and BOC-FPDNet for the urban and dense urban city environments within the LMMC. Analogous to the results in Fig. 6, data-driven DL-based approaches show a decline in error with increasing C/N_0 . They also demonstrate similar performance in both urban and dense urban settings, which is attributable to the fact that the performance of neural networks is primarily grounded in experiences learned from the dataset.

Regarding the FPDNet in Fig. 7, unlike the results shown in [15], there is not a significant difference in performance outcomes between VGGNet and ResNet. This suggests that although FPDNet effectively recognizes the global feature

TABLE IV
EARTH-SPACE LAND MOBILE CHANNEL MODEL DENSE
URBAN, URBAN PARAMETER [50]

Parameter	Value (max, min, mean, sigma)
Car Maximum Speed [km/h]	50
Satellite Elevation [°]	30
Satellite Azimuth [°]	-45
Carrier Frequency [GHz]	1.57542
Antenna Height [m]	2
Road Width [m]	15
Building Width [m]	-, 20, 30, 10
Building Height (Dense urban) [m]	150, 50, 100, 30
Building Height (Urban) [m]	100, 10, 50, 30
Gap between Buildings [m]	-, 10, 30, 20
Building Gap Likelihood (Dense urban)	0.1
Building Gap Likelihood (Urban)	0.4
Tree Height (Dense urban) [m]	8
Tree Height (Urban) [m]	7
Tree Diameter (Dense urban) [m]	5
Tree Diameter (Urban) [m]	4
Tree Trunk Length [m]	2
Tree Trunk Diameter [m]	2
Tree Attenuation [dB/m]	1.1
Tree Distance [m]	-, -, 40, 5
Pole Height (Dense urban) [m]	10
Pole Height (Urban) [m]	9
Pole Diameter [m]	0.2
Pole Distance [m]	-, -, 25, 10

of the ACF output compared to other CNN-based methods, reducing the ambiguity of main peak is of paramount importance. When contrasting BOC-FPDNet with other DL-based techniques, it shows comparable performance at a high SPC (16SPC), but it notably excels in FPD performance at a relatively lower SPC (4SPC). As distinguishing between side-peak and main peak becomes challenging at lower SPCs, the fact that BOC-FPDNet, benefitting from DeepSC, receives an ACF output with reduced main peak ambiguity and accurately localizes peaks indicating paths in the ACF output, elucidates its robustness.

Additionally, we analyzed the performance variations due to the high dynamic configuration of the receiver. This configuration is subject to correlation output attenuation due to acceleration [51]. The relative equivalent accelerations between the receiver and GNSS satellites were set at 5G and 10G (49m/s² and 98m/s²), respectively where G is referred to as the acceleration of gravity. The value of G is 9.8 m/s² on Earth. The elevation angle of the satellite from the receiver was set at 35°. Due to the variation in ACF distortion with T [51], we selected not only 1ms but also 10ms.

Fig. 8 shows the RMSE errors of FP delay estimation of BOC-FPDNet with different acceleration and T . The results show that the RMSE is slightly higher at $T = 10$ ms for the same SPC and acceleration. Regardless of the different T , there is no significant difference in RMSE within a high dynamic configuration. ACF output becomes negligibly distorted with extended T in the configuration. Consequently, this result indicates that BOC-FPDNet maintains great FPD performance even at short T , thereby ensuring its robustness in high dynamic configurations.

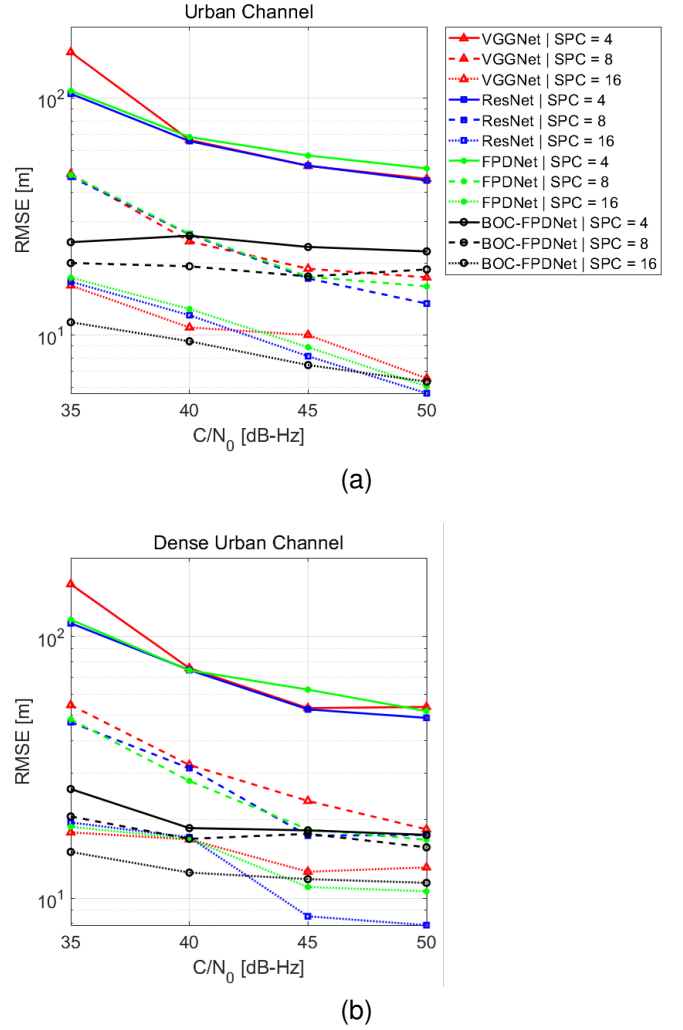


Fig. 7. RMS error of Urban and Dense Urban Channels.

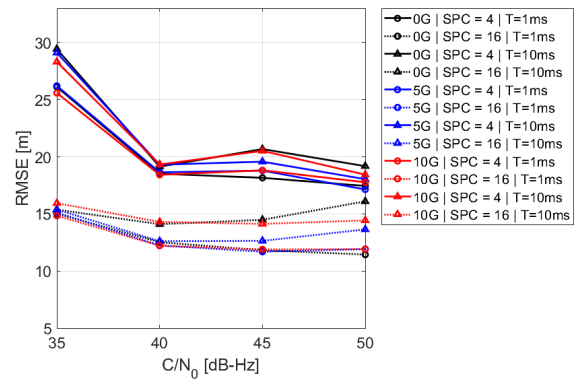


Fig. 8. RMS error of BOC-FPDNet according to the acceleration between receiver and satellite.

C. Positioning Improvement in Real Environments

To validate the enhanced positioning performance in real multipath channel environments due to the implementation of BOC-FPDNet, we utilized the TEX-CUP Dataset [52] released by the Radionavigation Lab at The University of Texas at Austin. The [52] dataset comprises data collected in

urban settings using multiple GNSS and Inertial Measurement Units (IMU). It also provides GNSS data from the base station required for Real-time Kinematic (RTK) processing. We targeted the BOC(1,1)-modulated signal of the GPS L1C. The dataset provided in [52] was collected using the low-cost Multi-GNSS Software Defined Radio (SDR) NTLab B1065U-12-X. Table V provides details about the channels used in the dataset.

TABLE V
CHANNEL DESCRIPTION OF TEX-CUP DATASET (GPS L1C) [52]

Parameter	Value
Sampling Frequency [MHz]	10.0
Spectral Inversion	False
Passband [MHz]	1573.32-1577.52
GPS L1C IF [MHz]	2.50

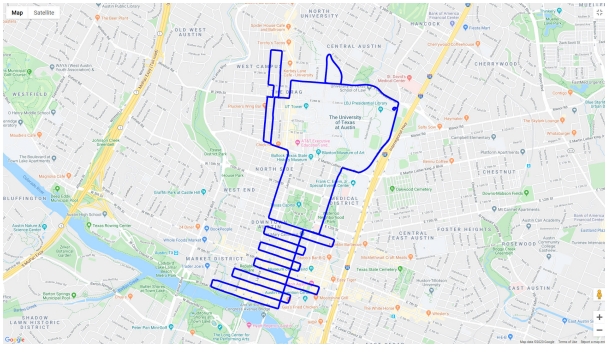


Fig. 9. Test route of TEX-CUP Dataset [52].

In the experiment, we received signals from five satellites. Only two of these signals passed through the same channels used in [15]. Unlike [15] which just takes a snapshot, our method tracks signals considering changing conditions. For time delay estimation, we use pilot signals and make feedback corrections, considering narrow band processing, where the pilot signal is exclusively limited to the BOC(1,1) of TM-BOC(6,1,1/11) modulation in GPS L1C. For the data signals, we use BOC-FPDNet and an Early-Late (EL) discriminator to estimate the delay. However, these errors aren't corrected or fed back. We consider EL discriminator as the conventional method. Similarly, BOC-FPDNet also corrects errors estimated from pilot signals, but for data signals, it only estimates the delay without corrections. The integration time is set to 10ms, with the focus on comparing multipath mitigation rather than robustness to noise.

Fig. 10 shows the distance root mean square (DRMS) errors and the root means square (RMS) errors, which represent the horizontal position error between reference positions and estimated positions to indicate accuracy, and the error between the estimated positions and the mean of estimated positions to indicate precision in our positioning techniques, respectively. As shown, our proposed BOC-FPDNet can reliably detect FP and helps improve positioning in multipath channel environment.

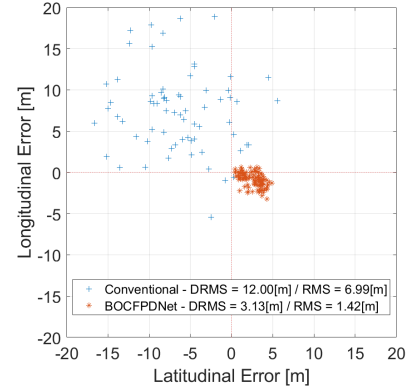


Fig. 10. Horizontal positioning error comparison.

V. CONCLUSION

In conclusion, this paper introduced the BOC-FPDNet, a two-stage architecture consisting of DeepSC and MHA-based MDE, designed to address the FP detection problem in multipath channel environments for BOC-modulated signals. From our investigations, the proposed method outperformed traditional mathematical models and other empirical neural network models such as VGGNet, ResNet, and the previously proposed FPDNet. Furthermore, the DeepSC component of the BOC-FPDNet offers advantages over the conventional SC techniques as it avoids the computationally intensive multi-subcorrelation and does not suffer from increasing noise interference. Through testing in real-world environments, we demonstrated that BOC-FPDNet significantly enhances positioning accuracy in multipath channels compared to conventional GNSS receivers. The proposed BOC-FPDNet can be presented as an effective solution for mitigating multipath errors in GNSS systems utilizing BOC-modulated signals. In future work, we plan to delve deeper into detecting the first arrival path in ACF outputs with more side peaks, such as Multiplexed binary offset carrier (CBOC, TMBOC), AltBOC.

VI. APPENDIX

In this section, we evaluate the Performance of DeepSC for BOC(6,1), which is a commonly used in BOC(kn, n) modulations. The number of side-peaks in BOC(kn, n) modulation is given by $4k - 2$. For instance, in the case of BOC(1,1) mainly discussed in this paper, where k equals 1, there are two side-peaks, and BOC(6,1) has 22 side-peaks. As noted in [5], with the increase in the order of k in BOC modulation, the width of the main peak decreases, enhancing range accuracy, but this also leads to an increase in the number of side-peaks, thereby increasing the ambiguity of the main peak. Therefore, as the order of k in BOC modulation increases, the importance of DeepSC is further emphasized.

The SC performance results of DeepSC for BOC(6,1) at 24SPC are presented in Table VI. Sim and NV in Table VI represent Similarity and Noise Variation, respectively. Despite the high SPC, BOC(6,1) shows a higher ambiguity in the main peak compared to BOC(1,1) in Table I, leading to a reduced performance. Across all C/N_0 s, DeepSC shows an

TABLE VI
SIDE-PEAKS CANCELLATION PERFORMANCE OF DEEPSC FOR BOC(6,1)

C/N ₀ (dB-Hz)	35		40		45		50	
	Sim (%)	NV (dB)	Sim (%)	NV (dB)	Sim (%)	NV (dB)	Sim (%)	NV (dB)
24SPC	83.21	2.20	84.31	1.91	85.10	1.76	85.61	1.53

average similarity and noise variation of 84.56% and 1.85dB, respectively. Similar to the trends shown in Table I, the performance decrease with increased noise. However, since a high level of similarity is maintained overall, the use of DeepSC is effective in multipath channel environments.

The side-peak cancelled ACF outputs of BOC(n,n) and BOC(m,n) modulation closely resemble the ACF output of BPSK modulation. Since MDE receives the side-peak cancelled ACF output as its input, BOC-FPDNet can be applied to all BOC modulation types.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1A2C3008370).

- [1] K. Borre, D. M. Akos, N. Bertelsen, P. Rinder, and S. H. Jensen, *A software-defined GPS and Galileo receiver: a single-frequency approach*. Springer Science & Business Media, 2007.
- [2] J. W. Betz and K. R. Kolodziejewski, "Generalized theory of code tracking with an early-late discriminator part ii: Noncoherent processing and numerical results," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1557–1564, 2009.
- [3] D. Torrieri, *Principles of spread-spectrum communication systems*. Springer, 2005, vol. 1.
- [4] C. Hegarty, J. W. Betz, and A. Saidi, "Binary coded symbol modulations for gnss," in *Proceedings of the 60th Annual Meeting of The Institute of Navigation (2004)*, 2004, pp. 56–64.
- [5] J. W. Betz, "Binary offset carrier modulations for radionavigation," *Navigation*, vol. 48, no. 4, pp. 227–246, 2001.
- [6] E. Gallon, M. Joerges, and B. Pervan, "Robust modeling of gnss orbit and clock error dynamics," *NAVIGATION: Journal of the Institute of Navigation*, vol. 69, no. 4, 2022.
- [7] Q. Zhao, J. Su, C. Xu, Y. Yao, X. Zhang, and J. Wu, "High-precision ztd model of altitude-related correction," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 609–621, 2022.
- [8] A. Steingass and A. Lehner, "Measuring the navigation multipath channel... a statistical analysis," in *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2004)*, 2004, pp. 1157–1164.
- [9] M.-D. Kim, J. Lee, J. Liang, and J. Kim, "Multipath channel characteristics for propagation between mobile terminals in urban street canyon environments," in *2015 17th International Conference on Advanced Communication Technology (ICACT)*. IEEE, 2015, pp. 511–516.
- [10] S.-H. Kong, "Toa and aod statistics for down link gaussian scatterer distribution model," *IEEE transactions on wireless communications*, vol. 8, no. 5, pp. 2609–2617, 2009.
- [11] Y. Lee, Y. Hwang, J. Y. Ahn, J. Seo, and B. Park, "Seamless accurate positioning in deep urban area based on mode switching between dgns and multipath mitigation positioning," *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [12] Y. Lee and B. Park, "Nonlinear regression-based gnss multipath modelling in deep urban area," *Mathematics*, vol. 10, no. 3, p. 412, 2022.
- [13] G. Zhang, L. Icking, L.-T. Hsu, and S. Schön, "A study on multipath spatial correlation for gnss collaborative positioning," in *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, 2021, pp. 2430–2444.
- [14] J. Lesouple, T. Robert, M. Sahmoudi, J.-Y. Tourneret, and W. Vigneau, "Multipath mitigation for gnss positioning in an urban environment using sparse estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1316–1328, 2018.
- [15] S.-H. Kong, S. Cho, and E. Kim, "Gps first path detection network based on mlp-mixers," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7764–7777, 2022.
- [16] E. S. Lohan, D. A. De Diego, J. A. Lopez-Salcedo, G. Seco-Granados, P. Boto, and P. Fernandes, "Unambiguous techniques modernized gnss signals: surveying the solutions," *IEEE Signal Processing Magazine*, vol. 34, no. 5, pp. 38–52, 2017.
- [17] K. Rouabah, S. Atia, M. Flissi, M. Salim Bouhlef, and S. Mezaache, "Efficient technique for dll s-curve side zero-crossings cancellation in global positioning system/galileo receiver," *IET Signal Processing*, vol. 13, no. 3, pp. 338–347, 2019.
- [18] P. Fine and W. Wilson, "Tracking algorithm for gps offset carrier signals," in *Proceedings of the 1999 national technical meeting of The Institute of Navigation*, 1999, pp. 671–676.
- [19] N. Martin, V. Leblond, G. Guillotel, and V. Heiries, "Boc (x, y) signal acquisition techniques and performances," in *Proceedings of the 16th international technical meeting of the satellite division of the institute of navigation (ION GPS/GNSS 2003)*, 2003, pp. 188–198.
- [20] O. Julien, C. Macabiau, M. E. Cannon, and G. Lachapelle, "Aspect: Unambiguous sine-boc (n, n) acquisition/tracking technique for navigation applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 150–162, 2007.
- [21] A. Burian, E. S. Lohan, and M. K. Renfors, "Efficient delay tracking methods with sidelobe cancellation for boc-modulated signals," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, pp. 1–20, 2007.
- [22] S. Kim, D. Chong, and S. Yoon, "A new gnss synchronization scheme," in *VTC Spring 2009-IEEE 69th Vehicular Technology Conference*. IEEE, 2009, pp. 1–5.
- [23] V. Heiries, D. Roviras, L. Ries, and V. Calmettes, "Analysis of non ambiguous boc signal acquisition performance," in *Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2004)*, 2004, pp. 2611–2622.
- [24] F. Wang, H. Gong, G. Liu, M. Li, C. Yan, T. Xia, X. Li, and J. Zeng, "DeepPicker: A deep learning approach for fully automated particle picking in cryo-em," *Journal of structural biology*, vol. 195, no. 3, pp. 325–336, 2016.
- [25] L. Garin and J.-M. Rousseau, "Enhanced strobe correlator multipath rejection for code & carrier," in *Proceedings of the 10th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1997)*, 1997, pp. 559–568.
- [26] B. Townsend and P. Fenton, "A practical approach to the reduction of pseudorange multipath errors in a ll gps receiver," in *Proceedings of the 7th International Technical Meeting of the Satellite Division of the Institute of Navigation, Salt Lake City, UT, USA*, 1994, pp. 20–23.
- [27] G. A. McGraw and M. S. Braasch, "Gnss multipath mitigation using gated and high resolution correlator concepts," in *Proceedings of the 1999 national technical meeting of the institute of navigation*, 1999, pp. 333–342.
- [28] F. Antreich, O. Esbrý-Rodríguez, J. A. Nossek, and W. Utschick, "Estimation of synchronization parameters using sage in a gnss-receiver," in *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005)*, 2005, pp. 2124–2131.
- [29] W. Nam and S.-H. Kong, "Least-squares-based iterative multipath super-resolution technique," *IEEE Transactions on signal processing*, vol. 61, no. 3, pp. 519–529, 2012.
- [30] D. Min, M. Kim, J. Lee, M. S. Ciriuc, M. Meurer, and J. Lee, "Dnn-based approach to mitigate multipath errors of differential gnss reference stations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 25 047–25 053, 2022.
- [31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.
- [32] C. Sun, A. Shrivastava, S. Singh, and A. Gupta, "Revisiting unreasonable effectiveness of data in deep learning era," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 843–852.
- [33] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 2014, pp. 740–755.
- [34] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *2012 IEEE conference on computer vision and pattern recognition*. IEEE, 2012, pp. 3354–3361.

- [35] S. A. Mohamed, M.-H. Hagbayan, T. Westerlund, J. Heikkonen, H. Tenhunen, and J. Plosila, "A survey on odometry for autonomous navigation systems," *IEEE access*, vol. 7, pp. 97 466–97 486, 2019.
- [36] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [37] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.
- [38] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.
- [39] E. S. Lohan, D. A. de Diego, J. A. Lopez-Salcedo, G. Seco-Granados, P. Boto, and P. Fernandes, "Unambiguous techniques modernized GNSS signals: Surveying the solutions," *IEEE Signal Process. Mag.*, vol. 34, no. 5, pp. 38–52, 2017. [Online]. Available: <https://doi.org/10.1109/MSP.2017.2711778>
- [40] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.
- [41] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.
- [42] S. Ruder, "An overview of multi-task learning in deep neural networks," *arXiv preprint arXiv:1706.05098*, 2017.
- [43] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [44] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [45] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *3rd International Conference on Learning Representations, ICLR 2015*, 2015.
- [46] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9268–9277.
- [47] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings*, 2011, pp. 315–323.
- [48] K. Yu, I. Sharp, and Y. J. Guo, *Ground-based wireless positioning*. John Wiley & Sons, 2009.
- [49] S.-H. Kong, "Statistical analysis of urban gps multipaths and pseudo-range measurement errors," *IEEE transactions on aerospace and electronic systems*, vol. 47, no. 2, pp. 1101–1113, 2011.
- [50] I. Recommendation, "Propagation data required for the design of earth-space land mobile telecommunication systems," *International Telecommunication Union: Geneva, Switzerland*, 2009.
- [51] Y. Luo, C. Yu, S. Chen, J. Li, H. Ruan, and N. El-Sheimy, "A novel doppler rate estimator based on fractional fourier transform for high-dynamic gnss signal," *IEEE Access*, vol. 7, pp. 29 575–29 596, 2019.
- [52] L. Narula, D. M. LaChapelle, M. J. Murrian, J. M. Wooten, T. E. Humphreys, E. de Toldi, G. Morvant, and J.-B. Lacambre, "Tex-cup: the university of texas challenge for urban positioning," in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 2020, pp. 277–284.



Dong-Hee Paek received the B.S. degree in robotics from Kwang-Woon University, South Korea, in 2019, and the M.S. degree from The Robotics Program at the Korea Advanced Institute of Science and Technology (KAIST) in 2021. He is currently pursuing the Ph.D. degree in the CCS Graduate School of Mobility from KAIST. His research interests include 4D Radar, AI-based perception, and autonomous vehicles.



Hong-Woo Seok received a B.S. degree in Geoinformatic Engineering from Inha University, Korea. He is currently pursuing the M.S. degree from Korea Advanced Institute of Science and Technology (KAIST). His research interests currently include GNSS signal processing, GNSS precise positioning, and multi-sensor navigation.



Seung-Hyun Kong (M'06–SM'16) is an Associate Professor in the CCS Graduate School of Mobility of Korea Advanced Institute of Science and Technology (KAIST), where he has been a faculty member since 2010. He received a B.S. degree in Electronics Engineering from Sogang University, Seoul, Korea, in 1992, an M.S. degree in Electrical and Computer Engineering from Polytechnic University (merged to NYU), New York, in 1994, a Ph.D. degree in Aeronautics and Astronautics from Stanford University, Palo Alto, in 2005. From 1997 to 2004 and from 2006 to 2010, he was with companies including Samsung Electronics (Telecommunication Research Center), Korea, and Qualcomm (Corporate R&D Department), San Diego, USA for advanced technology R&D in mobile communication systems, wireless positioning, and assisted GNSS. Since he joined KAIST as a faculty member in 2010, he has been working on various R&D projects in advanced intelligent transportation systems, such as robust GNSS-based navigation for urban environment, deep learning and reinforcement learning algorithms for autonomous vehicles, sensor fusion, and vehicular communication systems (V2X). He has authored more than 100 papers in peer-reviewed journals and conference proceedings and 12 patents, and his research group won the President Award (of Korea) in the 2018 international student autonomous driving competition host by the Korean government. He has served as an associate editor of IEEE T-ITS and IEEE Access, an editor of IET-RSN and the lead guest editor of the IEEE T-ITS special issue on "ITS empowered by AI technologie" and the IEEE Access special section on "GNSS, Localization, and Navigation Technologies". He has served as the program chair of IPNT from 2017 to 2019 in Korea, as a program co-chair of IEEE ITSC2019, New Zealand, and general chair of IEEE IV 2024, Korea.



Sangjae Cho received the B.S. degree in Department of Energy and Electrical Engineering from Korea Polytechnic University, Korea, in 2018 and M.S. degree in the CCS Graduate School of Mobility from Korea Advanced Institute of Science and Technology (KAIST), where he is currently pursuing the Ph.D. degree. His research interests include GNSS, Signal processing, Deep learning, Autonomous vehicle, and Wireless communication.