

# Fisheye Camera Aided GNSS NLOS Detection and Learning-based Pseudorange Bias Correction for Intelligent Vehicles in Urban Canyons

Runzhi Hu, Weisong Wen\*, and Li-Ta Hsu

**Abstract**— The global navigation satellite system (GNSS) is low-cost and is highly expected by intelligent vehicles, which can provide reliable absolute positioning service in open areas. Unfortunately, the performance of the GNSS positioning is significantly degraded in urban canyons, due to the notorious non-line-of-sight (NLOS) receptions and multipath interference caused by signal reflections. To fill this gap, this paper proposed a fisheye camera-based GNSS NLOS detection method, where the state-of-the-art *Swin Transformer* model is employed to segment the sky and non-sky view areas within the image. Instead of directly excluding the detected GNSS NLOS receptions that would significantly degrade the geometry of satellite distribution, this paper proposed to directly estimate the bias of the NLOS receptions and multipath by a tightly-coupled integration model of GNSS least square and the deep neural network (DNN). The effectiveness of the proposed method was validated on multiple datasets from urban canyons in Hong Kong. In particular, a GNSS NLOS detection accuracy of more than 99.0% is achieved in the evaluated dataset. By correcting the bias involved from the GNSS NLOS and multipath, improved positioning accuracy about 5% is obtained using the proposed method. To benefit the research community, we open-source a library, the *pyRTKLIB*, which can serve as an important bridge between the DNN and the conventional GNSS processing tools, the *RTKLIB*.

## I. INTRODUCTION

Global navigation satellite system (GNSS) provides convenient and low-cost global positioning for the increasing number of intelligent vehicles (IV) [1]. Centimeter-level positioning accuracy [2, 3] can be achieved by the GNSS positioning in open areas, with the help of the corrections from the reference stations. Unfortunately, most of these intelligent vehicles are operating in urban scenarios, where tall buildings reflect and obstruct the GNSS signals, referred to as urban canyons. The prevalence of urban canyons induces both multipath and non-line-of-sight (NLOS) bias, thereby exerting a substantial influence on the accuracy of pseudoranges [4, 5], which play a key role in GNSS measurements and reduce the precision of the GNSS positioning. According to a recent paper review [6], the GNSS NLOS reception is currently the dominant factor limiting the GNSS positioning accuracy in urban canyons. Addressing this issue constitutes an immediate imperative.

*Receiver autonomous integrity monitoring (RAIM) and signal-to-noise ratio (SNR) based GNSS NLOS mitigation:* To solve this problem, it is intuitive to detect and exclude the GNSS NLOS satellites. The RAIM [7] is a technique

developed to guarantee navigation integrity through self-contained fault detection at the receiver. RAIM can detect faulty measurements, which include multipath and NLOS. Walter [8] tested the validity of this assumed hypothesis using Chi-square distributed sum-of-squared residuals. However, these RAIM methods relied on the redundancy of the healthy measurements, with the assumption that the healthy measurements are more than the polluted measurements. Unfortunately, this assumption can be easily violated in urban canyons. Instead of relying on the measurements redundancy, the work in Uaratanawong [9] and Zhang [10] directly analyze the GNSS raw features, including the SNR and elevation angles to model the multipath error. However, only using GNSS raw measurements is hard to model the error in complex scenarios such as urban canyons.

*3D mapping/Light Detection And Ranging (LiDAR) aided GNSS NLOS mitigation:* Given the fact that the GNSS NLOS receptions are caused by the surrounding environments, the 3D building model is a direct resource to represent the environment. GNSS shadow matching [11] use a 3D building model to test the satellite visibility on the hypothesized positions. Range-based 3D mapping aided (3DMA) GNSS methods [12] aim to correct NLOS-affected measurements by using the ray-tracing method. However, the precision of the above 3DMA methods requires accurate 3D models, which are hard to access in some places. Meanwhile, another limitation of the 3DMA GNSS positioning method is its strong reliance on the initial guess of the receiver's location. Instead of relying on the prior 3D mapping information, our previous work continuously proposed to use the 3D LiDAR to detect the potential NLOS measurements [13] and even correct their errors [14]. With the help of 3D LiDAR, the environment can be reconstructed in real-time for further GNSS NLOS detection, which opens a new window for their mitigation in urban canyons. However, the high cost of the 3D LiDAR sensors limits their massive deployment.

*Vision-aided GNSS NLOS mitigation:* Using a sky-pointing camera to capture the sky view is a cheaper but reliable way to classify the visibility of satellites. In 2009, the pioneering work by Meguro [15] used an omnidirectional infrared (IR) camera to eliminate the NLOS satellites. The main idea is to use the sky-pointing camera to capture the sky-view image and then project the received satellite signals onto the latter to further identify the NLOS receptions. In their methods, the sky-pointing infrared camera could naturally and easily segment the sky and obstructions. However, the camera was also affected by other heated things such as the sun. Thanks to the advancement in computer vision (CV) advanced, color-based segmentation became

All authors are with the Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong. (The corresponding author to provide e-mail: welson.wen@polyu.edu.hk).

possible. Yao [16] directly used the grayscale threshold to filter the sky. Suzuki [17] used a mean shift algorithm [18] to cluster the sky. Our previous work in [19] used the median filter to segment the sky. However, these conventional filtering methods are sensitive to clouds and light changes, which are not robust enough to classify the sky and non-sky area. *In short, a reliable and accurate sky and non-sky segmentation method is highly expected for robust vision-based NLOS detection.* As neural networks demonstrated their potential in CV, a large number of learning-based methods came out. In 2017, a Transformer structure [20] was proposed in the natural language processing (NLP) field. It was subsequently utilized in computer vision [21], referred to as Vision Transformer (ViT), and demonstrated comparable performance to convolutional neural networks (CNNs). Recently, Swin Transformer [22] reached the state-of-the-art. In contrast to ViT, Swin Transformer used shifted windows to reduce the computation and achieved superior performance. *Can we use this advancement to improve vision-based GNSS NLOS detection?* This is the first question this paper would answer. In fact, the NLOS exclusion would lead to poor satellite geometry in urban canyons, as only the satellites received from the high elevation angles survive to the NLOS exclusion. As a result, the direct NLOS exclusion could potentially lead to degraded performance and this phenomenon has been consistently demonstrated in our prior research in [14, 23, 24]. Given the actively perceived information from the sky-pointing camera and the passively received GNSS features, such as the SNR and elevation angles, *can we also adopt the DNN to correct the bias involved in the NLOS and multipath?* This is the second question this paper would answer. Thus, in this work, our contributions are as follows:

1) *Swin Transformer aided GNSS NLOS detection with the fisheye camera:* Our work proposes an NLOS detection method using the Swin Transformer to segment fisheye image. This process significantly improves the accuracy of the NLOS detection.

2) *DNN-based GNSS bias correction:* A deep neural network architecture is proposed to correct the pseudorange bias involved by the detected NLOS and multipath receptions. This method significantly improves the GNSS positioning accuracy.

3) *Extensive experimental validation:* We test our proposed method on four individual datasets, including light urban and deep urban scenarios. Meanwhile, we also open source the PyRTKLIB at <https://github.com/IPNL-POLYU/pyrtklib>, which is a Python binding for RTKLIB [2]. In particular, the RTKLIB is a very famous and popular library in the GNSS research field for GNSS data processing. The developed PyRTKLIB effectively bridges the gap between the RTKLIB and the state-of-the-art machine learning networks. We believe that the PyRTKLIB would lead to a positive impact on both the academic and industry fields.

This paper is structured as followed: The proposed method is presented in Section II and the detail is elaborated in Section III. The experiments and discussion are conducted in Section IV. Conclusions and future work are given in Section V.

## II. OVERVIEW OF THE PROPOSED METHOD

The framework of the proposed method is shown in Fig. 1. The system is divided into two parts: Swin Transformer-aided NLOS detection and pseudorange bias correction. The input of the system includes the raw GNSS measurements from the user and the sky view image captured by the sky-pointing camera. Note that the heading information of the vehicle is provided by the electrical compass sensor. These data are first synchronized based on timestamps and the time difference is less than 0.03 seconds. Note that the corrections from the reference stations are used to remove the systematic clock and atmospheric errors involved in the pseudorange measurements. First, based on the least square estimation using differential GNSS (DGNSS), the satellite elevation and azimuth angles can be inferred from the receiver measurements. Therefore, the satellites can be projected onto the fisheye image plane using the principles of fisheye camera imaging. The NLOS satellites can be detected accordingly. Since the pre-trained Swin Transformer model is trained on the ADE20K dataset [25], which consists of non-fisheye images, the fisheye images need to be expanded into panoramic images. The expanded images are then passed through the Swin Transformer, generating a sky mask that accurately identifies the sky region. This enables straightforward NLOS satellite detection based on the precise sky mask. Second, features such as double-differenced residuals, SNR, elevation angle, LOS/NLOS flag, SNR of the master satellite, which is designated as the reference point from which single differences are calculated for measurements obtained from other satellites, and LOS/NLOS flag of the master satellite are extracted from each satellite in a positioning epoch. These features are then fed into the deep neural network (DNN) layer to predict the pseudorange bias. Finally, the predicted biases are added to the raw pseudorange measurements, and another round of DGNSS is performed to obtain the improved position, based on the corrected pseudorange measurements.

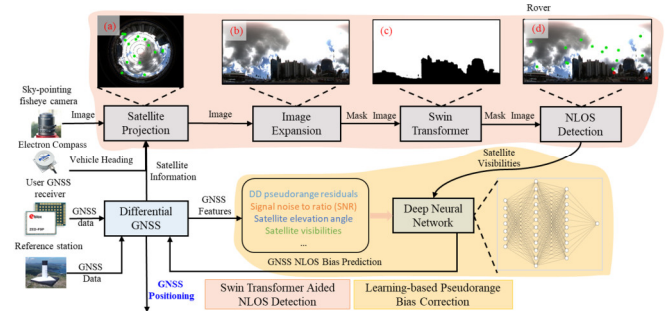


Fig. 1 The framework of the proposed method in this paper.

## III. METHODOLOGY

The detail of the proposed methodology is presented in this section.

### A. Fisheye Camera Aided GNSS NLOS Detection

As shown in Fig. 1, the first step in the process is to project the satellites onto the fisheye image. This requires synchronizing the data from three sensors: (1) the GNSS receiver, which provides satellite position including azimuth and elevation angles; (2) the fisheye camera, which provides the skyward image; (3) the electrical compass, which

provides the azimuth angle of the vehicle. To provide a clear explanation of the projection process, it is necessary to introduce the principle of fisheye camera imaging.

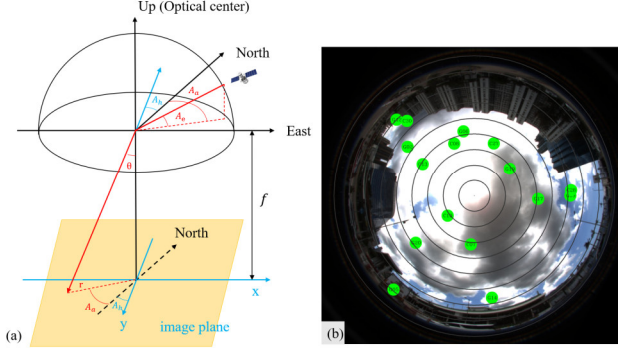


Fig. 2 (a) Illustration of the fisheye camera model and (b) illustration of the satellite projection to the image.

As depicted in Fig. 2 (a), the main coordinate is in East-North-Up (ENU) frame. The red line is the incident line of sight of the satellite.  $A_a$  and  $A_e$  correspond to the azimuth angle and elevation angle of the satellite, respectively.  $A_h$  is the heading angle, of the vehicle.  $\theta$  is the refraction angle.  $r$  is the radius of the satellite projection.  $f$  is the focal length. The fisheye camera model describes the relationship between the incident angle  $\alpha$  and the refraction angle  $\theta$ , where  $\alpha = \frac{\pi}{2} - A_e$ . There are several models, including equidistance projection, orthogonal projection, equisolid angle projection, and stereographic projection. All of these models can be expressed using the Taylor Expansion with respect to the incident angle  $\alpha$ :

$$\tan(\theta) = \alpha + \sum_{i=1}^n k_i \alpha^{2i+1} \quad (1)$$

$$r = f \times \tan(\theta) \quad (2)$$

where  $k_i$  denotes the  $i$ th order distortion parameter associated with the fisheye camera and can be calibrated by capturing a chessboard image. The fisheye camera utilized in our system adheres to the equidistant projection rule, implying that all the  $k$  is zero. And with the satellite azimuth angle and vehicle heading angle, the satellite coordinate in the image plane can be determined by:

$$\begin{cases} x = r \times \sin(A_a + A_h) \\ y = r \times \cos(A_a + A_h) \end{cases} \quad (3)$$

Fig. 2(b) illustrates the projected satellites in the fisheye image. Subsequently, the fisheye image should be expanded into a panorama image, since the pre-trained model is not specifically trained on fisheye images and may not segment them accurately. The expansion follows below equations:

$$\begin{cases} u' = r + \frac{u}{2} \cos\left(2\pi \frac{v}{d}\right) \\ v' = r + \frac{u}{2} \sin\left(2\pi \frac{v}{d}\right) \end{cases} \quad (4)$$

where  $d$  is the diameter of the fisheye image, and  $u'$  and  $v'$  represent the pixel coordinate in the fisheye image frame, while  $u$  and  $v$  denote the pixel coordinates after the expansion. The coordinate of the satellites in the fisheye

image is also transformed using these equations and Fig. 1(b) can be obtained.

The panorama image is then fed as input to the Swin Transformer for segmentation. An example of the image segmentation result, so-called the sky mask, is shown in the Fig. 1(c). By utilizing the generated sky mask, it becomes straightforward to determine whether the satellites are obstructed or not by examining their position relative to the sky region. The detection result is displayed in Fig. 1(d), with LOS and NLOS marked in green and red, respectively.

### B. Learning-based Pseudorange Bias Correction

The second part focuses on the learning-based pseudorange bias correction. In urban canyons, the accuracy of pseudorange measurements is compromised due to multipath and NLOS effects. Physical simulation of signal propagation using ray-tracing with a 3D building model can address this issue. However, the availability of accurate 3D building models is limited in most cities. *Alternatively, if the environmental effects can be considered as a function, the GNSS signals, including azimuth angle, elevation angle, and SNR, serve as the inputs, while the observation measurements, such as SNR, pseudorange, along with the pseudorange bias, act as the outputs. This implies the presence of a hidden model connecting the inputs, outputs, and pseudorange bias.*

To capture this hidden model, a DNN model is proposed to learn and predict the pseudorange bias using 7 features. Before delving into these features, it is necessary to introduce the DGNSS method to provide a clear understanding.

#### 1) DGNSS Solution

DGNSS is a method using a reference station to eliminate the troposphere and ionospheric delay. A GNSS measurement can be modeled as follow:

$$p_r^s = r_r^s + c\Delta t + I_r^s + T_r^s + \epsilon_{r,NLOS/MP}^s + \epsilon_r^s \quad (5)$$

$$r_r^s = \sqrt{(x^s - x_r)^2 + (y^s - y_r)^2 + (z^s - z_r)^2} \quad (6)$$

where the subscript  $r$  represents the receiver, and superscript  $s$  represents the satellite index.  $p$  is the pseudorange measured by receiver,  $r_r^s$  is the geometric distance between the satellite and receiver,  $c$  is the speed of light and  $\Delta t$  is the clock error.  $x^s, y^s, z^s, x_r, y_r, z_r$  are the coordinate of satellite and receiver in Earth-Centered, Earth-Fixed (ECEF) coordinate system.  $I$  and  $T$  are the ionospheric and troposphere delay.  $\epsilon_{NLOS/MP}$  is the error caused by NLOS or multipath. The last  $\epsilon$  denotes the Gaussian noise. As shown in Fig. 3, DGNSS relies on a nearby reference station under open sky without NLOS and multipath, whose position is known accurately. According to [26], if the distance between the station and receiver is less than 10km, the atmospheric conditions can be regarded as the same. Then, ionospheric and troposphere delay can be eliminated by pseudorange double difference:

$$\begin{aligned} (r_{r_1}^s - r_{r_1}^{s_m}) - (r_{r_2}^s - r_{r_2}^{s_m}) &= (p_{r_1}^s - p_{r_1}^{s_m}) - (p_{r_2}^s - p_{r_2}^{s_m}) \\ r_{r_1}^s - r_{r_1}^{s_m} &= (p_{r_1}^s - p_{r_1}^{s_m}) - (p_{r_2}^s - p_{r_2}^{s_m}) + (r_{r_2}^s - r_{r_2}^{s_m}) \end{aligned} \quad (7)$$

where  $s_m$  denotes the main satellite in single difference and subscripts  $r_1$  and  $r_2$  denote the receiver and reference station

respectively. The geometric distance between the satellite and reference station, denoted by  $r_2$  including  $x_{r_2}, y_{r_2}, z_{r_2}$  is known with accuracy. It is important to note that NLOS and multipath errors are not considered here. In equation (7), there are three unknowns:  $x_{r_1}, y_{r_1}, z_{r_1}$ . Therefore, 3 or more equations are required to solve these unknowns. To address this, gradient descent and least square are used. The observation function  $h(\mathbf{X})$  and observation measurements  $\mathbf{Z}$  can be formulated as follows:

$$\mathbf{Z} = h(\mathbf{X}) = \begin{bmatrix} r_{r_1}^1 - r_{r_1}^{s_m} \\ r_{r_1}^2 - r_{r_1}^{s_m} \\ \dots \\ r_{r_1}^n - r_{r_1}^{s_m} \end{bmatrix}, \mathbf{X} = (x_{r_1}, y_{r_1}, z_{r_1}) \quad (8)$$

$$\mathbf{Z}' = \begin{bmatrix} (p_{r_1}^1 - p_{r_1}^{s_m}) - (p_{r_2}^1 - p_{r_2}^{s_m}) + (r_{r_2}^1 - r_{r_2}^{s_m}) \\ (p_{r_1}^2 - p_{r_1}^{s_m}) - (p_{r_2}^2 - p_{r_2}^{s_m}) + (r_{r_2}^2 - r_{r_2}^{s_m}) \\ \dots \\ (p_{r_1}^n - p_{r_1}^{s_m}) - (p_{r_2}^n - p_{r_2}^{s_m}) + (r_{r_2}^n - r_{r_2}^{s_m}) \end{bmatrix} \quad (9)$$

Then the equation (7) can be formulated in matrix form by  $\mathbf{Z}'$  and  $\mathbf{Z}$  as:

$$\mathbf{Z} = \mathbf{Z}' \quad (10)$$

And the Jacobian matrix of the function is:

$$\mathbf{H}(\mathbf{X}) = \begin{bmatrix} -\frac{x^1 - x_{r_1}}{r_{r_1}^1} - \left(-\frac{x^{s_m} - x_{r_1}}{r_{r_1}^{s_m}}\right) & -\frac{y^1 - y_{r_1}}{r_{r_1}^1} - \left(-\frac{y^{s_m} - y_{r_1}}{r_{r_1}^{s_m}}\right) & -\frac{z^1 - z_{r_1}}{r_{r_1}^1} - \left(-\frac{z^{s_m} - z_{r_1}}{r_{r_1}^{s_m}}\right) \\ -\frac{x^2 - x_{r_1}}{r_{r_1}^2} - \left(-\frac{x^{s_m} - x_{r_1}}{r_{r_1}^{s_m}}\right) & -\frac{y^2 - y_{r_1}}{r_{r_1}^2} - \left(-\frac{y^{s_m} - y_{r_1}}{r_{r_1}^{s_m}}\right) & -\frac{z^2 - z_{r_1}}{r_{r_1}^2} - \left(-\frac{z^{s_m} - z_{r_1}}{r_{r_1}^{s_m}}\right) \\ \dots & \dots & \dots \\ -\frac{x^n - x_{r_1}}{r_{r_1}^n} - \left(-\frac{x^{s_m} - x_{r_1}}{r_{r_1}^{s_m}}\right) & -\frac{y^n - y_{r_1}}{r_{r_1}^n} - \left(-\frac{y^{s_m} - y_{r_1}}{r_{r_1}^{s_m}}\right) & -\frac{z^n - z_{r_1}}{r_{r_1}^n} - \left(-\frac{z^{s_m} - z_{r_1}}{r_{r_1}^{s_m}}\right) \end{bmatrix} \quad (11)$$

When given a position  $\mathbf{X}_i$ , the value of the observation function is  $\mathbf{Z}_i = h(\mathbf{X}_i)$ . Then gradient descent and least square are used to update the position iteratively:

$$\mathbf{Z}' - \mathbf{Z}_i = \mathbf{H}(\mathbf{X}_i) \Delta \mathbf{x} \quad (12)$$

$$\Delta \mathbf{x} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T (\mathbf{Z} - \mathbf{Z}_i) \quad (13)$$

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \Delta \mathbf{x} \quad (14)$$

where  $\mathbf{H}$  in equation (13) is the short for  $\mathbf{H}(\mathbf{X}_i)$  in equation (11). When  $\|\Delta \mathbf{x}\|$  is less than a predefined threshold or the iteration time exceeds a specified limit, the iteration process is stopped. In our work, we set the threshold to 0.01 meters and the time limit to 20 iterations, and the initial guess for the position is (0,0,0).

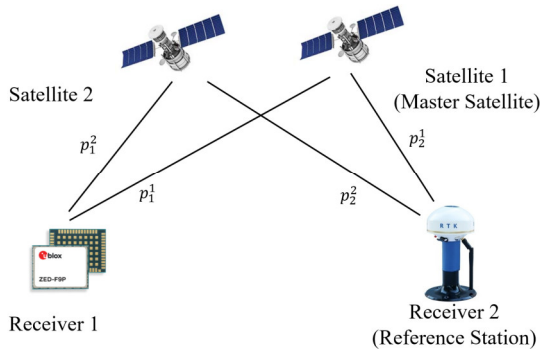


Fig. 3 A receiver and a reference station

## 2) Ground Truth Label

In equation (7), NLOS and multipath are ignored, and the time bias  $\Delta t$ , ionospheric delay  $I$  and troposphere delay  $T$  can

be eliminated. However, if NLOS and multipath exist, the equation (10) should be:

$$\mathbf{Z}' = \mathbf{Z} + \begin{bmatrix} \epsilon_{NLOS/MP}^1 - \epsilon_{NLOS/MP}^{s_m} \\ \epsilon_{NLOS/MP}^2 - \epsilon_{NLOS/MP}^{s_m} \\ \dots \\ \epsilon_{NLOS/MP}^n - \epsilon_{NLOS/MP}^{s_m} \end{bmatrix} \quad (15)$$

where  $\mathbf{Z}'$  is the real observation measurements including the NLOS and multipath error. The ground truth position  $\mathbf{X}_{gt}$  is provided by the real-time kinematic (RTK) GNSS/INS integrated solution from NovAtel SPAN-CPT, which can achieve centimeter-level accuracy in RTK mode[27]. The ground truth can provide accurate geometric distance  $r$ . Thus, the residual of the ground truth  $\mathbf{Z}_{gt} = h(\mathbf{X}_{gt})$  can be expressed as:

$$\mathbf{Y} = \mathbf{Z}' - \mathbf{Z}_{gt} = \begin{bmatrix} \epsilon_{NLOS/MP}^1 - \epsilon_{NLOS/MP}^{s_m} \\ \epsilon_{NLOS/MP}^2 - \epsilon_{NLOS/MP}^{s_m} \\ \dots \\ \epsilon_{NLOS/MP}^n - \epsilon_{NLOS/MP}^{s_m} \end{bmatrix} \quad (16)$$

$\mathbf{Y}$  is the label, which contains the NLOS and multipath error of the chosen satellite and the master satellite, which should be corrected.

## 3) Input Features Selection

### a) Double-differenced Residual

Double-differenced residual is a feature gotten from DGNSS solution. The double-differenced residual is formulated as:

$$\mathbf{DD}_{res} = \mathbf{Z}' - \hat{\mathbf{Z}} = \mathbf{Z}' - h(\hat{\mathbf{X}}) \quad (17)$$

$$dd_{res}^s = (p_{r_1}^s - p_{r_1}^{s_m}) - (p_{r_2}^s - p_{r_2}^{s_m}) - ((\hat{r}_{r_1}^s - r_{r_1}^{s_m}) - (\hat{r}_{r_2}^s - r_{r_2}^{s_m})) \quad (18)$$

where  $\hat{\mathbf{Z}}$  is the observation function value using the solution. In equation (17), if  $\hat{\mathbf{Z}}$  approximate  $\mathbf{Z}'$ , the position should be accurate. However, due to the presence of NLOS and multipath, the solution is likely to have an error greater than zero. This serves as a significant feature of the model.

### b) Elevation Angle

The elevation angle represents the height of the satellite. It is a crucial measurement that serves to assess reliability and aids in NLOS detection, as higher elevation angles typically indicate better visibility. In particular, the satellite with lower elevation angle tends to be reflected by surrounding buildings. However, since NLOS has already been detected in the initial step of our model, the elevation angle in our approach assists in modeling the hidden model which is described in III. B.

### c) Signal-to-noise Ratio (SNR)

The SNR represents the quality of the satellite signal and plays a crucial role in modeling NLOS and multipath effects. In particular, reflections tend to considerably reduce the SNR. Therefore, in our model, SNR is utilized as an input parameter.



#### d) LOS/NLOS Flag

The LOS/NLOS flag is outputted by the NLOS detection model. If the satellite is classified as a LOS satellite, then it should be 1. Conversely, if it is identified as a NLOS satellite, the flag is set to 0.

#### e) Elevation Angle, SNR and LOS/NLOS Flag of the Master Satellite

In the conventional approach, the satellite with the highest elevation angle is typically selected as the master satellite. However, in urban canyons, the satellite with a higher elevation angle may be obstructed by tall buildings on either side, while a satellite with a lower elevation angle may have a direct line of sight through a narrow strip of the sky. Fortunately, our NLOS detection algorithm accurately identifies both NLOS and LOS satellites. In our work, we choose the LOS satellite with the highest elevation angle as the master satellite. If there are no LOS satellites available, we select the satellite with the highest elevation angle as the master satellite instead. Regardless of the case, according to equation (15), the pseudorange that needs correction includes the NLOS and multipath contributions from both the chosen satellite and the master satellite. Hence, information about the master satellite is also necessary for our model. The elevation angle, SNR, and LOS/NLOS flag of the master satellite are utilized as three additional features. The model aims to learn the combination of NLOS and multipath effects from both the chosen satellite and the master satellite. These aforementioned 7 features are used as inputs to the model. It is worth noting that both LOS and NLOS satellites are included in the network to predict the bias, as the model is also expected to capture the multipath effects from the LOS satellites.

#### 4) Proposed Network

The proposed network is a DNN designed to predict the pseudorange bias for each satellite within a single frame. The network takes seven features as input and generates the predicted pseudorange bias as output. The architecture of the DNN is depicted in Fig. 4.

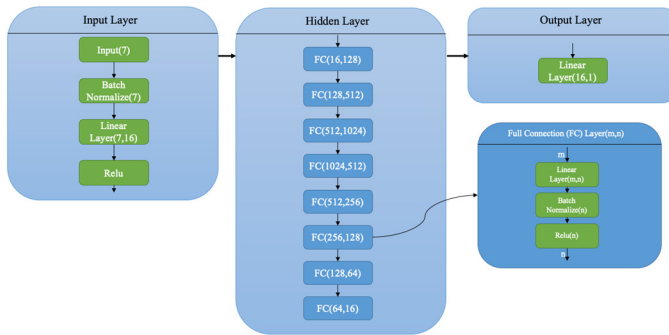


Fig. 4 Proposed DNN Architecture

The input is initially passed through a Batch Normalization (BN) layer for normalization. The normalized result then proceeds through a Linear Layer, increasing the dimension to 16, with the Rectified Linear Unit (ReLU) function employed as the activation function. Following this, the output passes through eight hidden layers, each consisting of a Linear Layer to control the dimension, a BN layer, and an activation layer. The dimensions are progressively

increased to 1024 and then gradually reduced to 16. Finally, a linear layer is utilized as the output layer. Given the depth of the network, the inclusion of BN layers is essential. Without BN layers, the gradient may vanish, and the network fails to converge.

## IV. EXPERIMENTS

### A. Experiment Setup

**Data collection platform and scenarios:** The proposed method is evaluated using four datasets, which were collected in urban areas of the Hong Kong Whampoa and Kowloon Tong (KLT) areas. More details of the data collection platform can be found by our recently opensource dataset [27]. In particular, the sky view image from the sky-pointing camera is collected at 30Hz. The raw GNSS measurements (GPS and Beidou) are collected at L1 frequency band using the u-blox GNSS receiver. Fig. 6 provides detailed information about the four datasets. Specifically, (a), (b), and (c) refer to the KLT1, KLT2, and KLT3 datasets, respectively, which were collected in a nearby region. The scenario in KLT is classified as a light urban scenario. On the other hand, scenario (d) in Whampoa represents a deep urban scenario characterized by significant presence of NLOS and multipath effects. The ground truth of positioning is provided by SPAN-CPT [27]. The four datasets consist of 203, 209, 404, and 1161 epochs, respectively.

**Training of the DNN for pseudorange bias prediction:** Regarding the training process, the model was trained on Ubuntu 20.04 using PyTorch. The system configuration included an Nvidia RTX 4090 graphics card, an AMD 5900x CPU, and 128 Gigabytes of memory. The model was trained for 2000 epochs using the KLT3 dataset and tested on the KLT1, KLT2, and Whampoa datasets. Mean Square Error (MSE) was utilized as the loss function, and Adam Optimizer [28] was employed as the optimizer with an initial learning rate of 0.01. The training process was repeated more than 5 times, and each model produced nearly identical results.

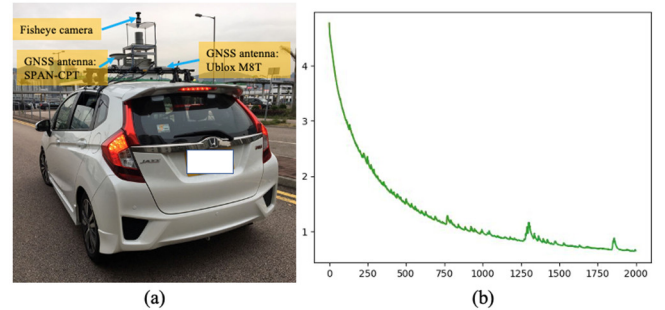
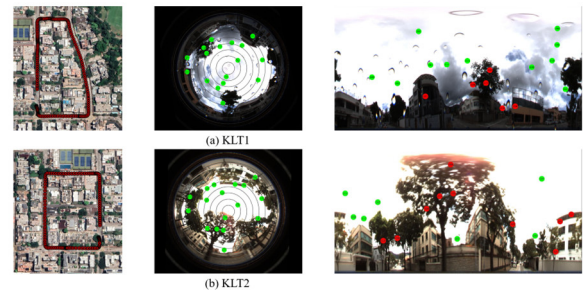


Fig. 5 (a) The experimental vehicle and sensor kit. (b) The training process.



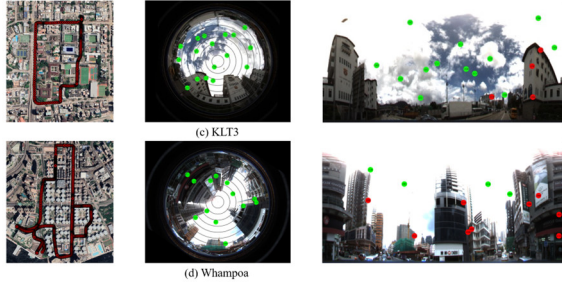


Fig. 6 The detail of the four datasets. In each row, the left image is the evaluated trajectory of the dataset, the middle image is a fisheye image with satellite sky plot, and the right image shows an example of the NLOS detection result.

### B. Result of NLOS Detection

The evaluation of the NLOS detection results involves assessing the accuracy of the NLOS detection. Firstly, the projected positions of the satellites are compared to the results obtained from the 3DMA method by our previous work [29], and the relative positions are found to be generally correct based on the 3DMA results. The compared results are shown at [https://github.com/ebhrz/NLOS\\_detection](https://github.com/ebhrz/NLOS_detection).

With the accurate projected satellite positions established, the LOS and NLOS satellites can be manually labeled on the image. If the pixel corresponding to the projected satellite position does not belong to the sky region, it is detected as an NLOS satellite. The evaluation metrics include accuracy, precision, and recall rate, which are summarized in Table 1.

Table 1. The evaluation result of the proposed NLOS detection method

Dataset		NLOS(P)	LOS(N)	Accuracy	Precision	Recall
KLT1	T	1408	1859	99.45%	99.09%	99.65%
	F	13	5			
KLT2	T	1706	1747	99.37%	99.12%	99.66%
	F	15	6			
KLT3	T	2262	3909	99.53%	99.23%	99.59%
	F	18	9			
Whampoa	T	7903	6316	99.02%	98.99%	99.38%
	F	81	49			

The detection results are presented in the table, where 'T' represents true, and 'F' represents false. In the evaluation, NLOS is considered positive, and LOS is considered negative. As indicated by the results, the proposed method achieves a high accuracy of 99% along with a high precision rate and recall rate. Fig. 8 showcases some of the cases where false detections occurred.

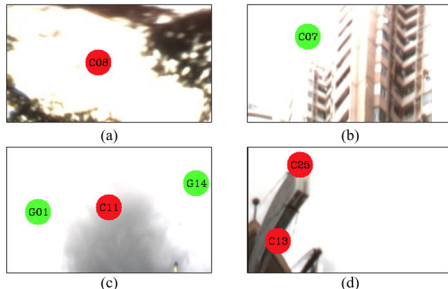


Fig. 8 False detection cases.

In Fig. 8 (a), (b), and (c), the false detections are attributed to inaccuracies in sky segmentation. The sky is not correctly recognized in (a), and the gray cloud is not identified as part of the sky. In (b), although C07 should be classified as NLOS, the long exposure time of the camera results in the bright image area being mistakenly segmented as sky. These errors can potentially be reduced by improving the accuracy of the semantic segmentation algorithm. Fig. 8 (d) represents a special case. In this image, the satellites appear to be blocked by a sign. However, considering the diffraction of the carrier wave, the carrier waves of the two satellites should not be blocked. Therefore, this type of error remains unsolved at present.

### C. Result of Pseudorange Correction

The evaluation of our method relies on the positioning results after pseudorange correction, which serves as a key metric. The model utilized for this evaluation is trained on the KLT3 dataset and tested on the KLT1, KLT2, and Whampoa datasets. For each satellite in each frame, the pseudorange bias is predicted and subsequently corrected by employing the following procedure:

$$b^s = D(\mathbf{I}^s), \mathbf{I}^s = (i_1^s, i_2^s, i_3^s, i_4^s, i_5^s, i_6^s, i_7^s) \quad (19)$$

$$z^{s''} = z^{s'} - b^s$$

$$= (p_{r_1}^s - p_{r_1}^{sm}) - (p_{r_2}^s - p_{r_2}^{sm}) + (r_{r_2}^s - r_{r_2}^{sm}) - b^s \quad (20)$$

where  $\mathbf{I}^s$  contains the seven input features of the  $s$ th satellite,  $D$  is the DNN model,  $b^s$  is the predicted bias of the  $s$ th satellite, and  $z^{s''}$  is the corrected observation value. Then the corrected observations in one frame is represented as:

$$\mathbf{Z}'' = \begin{bmatrix} z^{1''} \\ z^{2''} \\ z^{3''} \\ \vdots \\ z^{n''} \end{bmatrix} \quad (21)$$

Then the  $\mathbf{Z}'$  in equation (12) is changed to  $\mathbf{Z}''$ , and the new solution is the corrected solution.

Table 2. The evaluation of the bias corrected method

Dataset	Method	Min(m)	Max(m)	Mean(m)	Improvement
KLT1	RTK	0.21	7.03	2.74	4.74%
	DLS	0.25	10.69	2.79	6.45%
	DBCLS	0.10	17.66	2.61	/
KLT2	RTK	0.22	13.17	4.04	19.06%
	DLS	0.61	17.75	4.82	32.16%
	DBCLS	0.10	9.77	3.27	/
Whampoa	RTK	0.75	219.90	22.58	10.94%
	DLS	0.30	218.71	21.12	4.78%
	DBCLS	0.05	232.55	20.11	/

In this evaluation, three results are compared: the RTK solution obtained from RTKLIB, where the mode is kinematic, and the solution type is forward; the DGNSS least square (DLS) solution; and the DGNSS bias corrected least square (DBCLS) solution derived from our proposed method. The solutions are transformed to the East-North-Up (ENU) frame, and the evaluation focuses on the accuracy in the

eastern and northern directions. The metric used is the Euclidean distance between the solutions and the ground truth. It's important to note that all the solutions are for a single frame and have not been post-processed using a Kalman Filter. Table 2 and Fig. 9 present the evaluation results.

The Improvement column in Table 2 quantifies the enhancements achieved by our proposed method, DBCLS, in comparison to the original method.

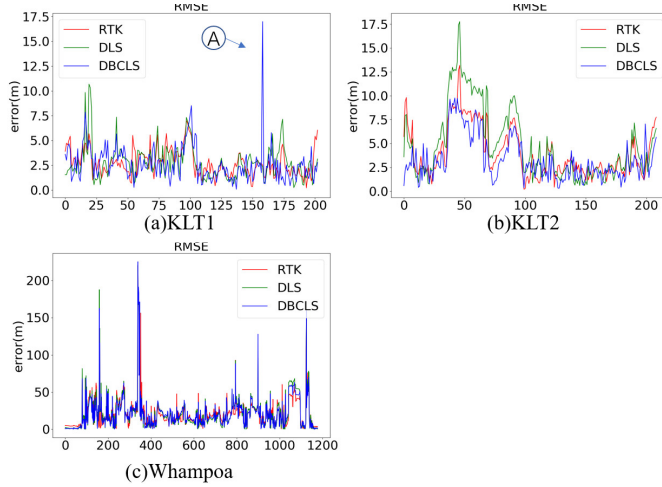


Fig. 9 The error of the above methods in three datasets. The red line is the error of RTK, the green line is the error of DLS, and the blue line is the error of DBCLS.

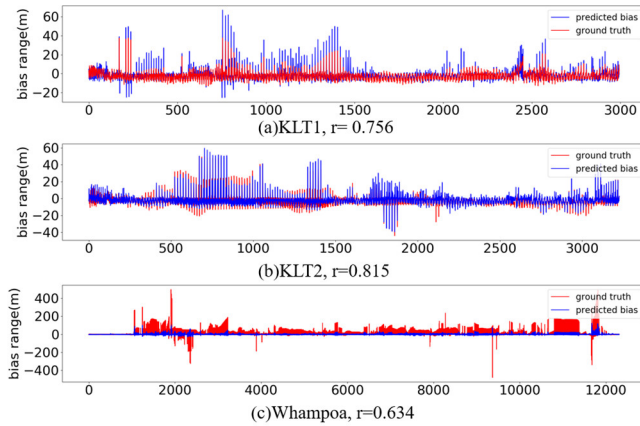


Fig. 10 The predicted bias, ground truth bias and their correlation coefficient of the three datasets

Based on the results, it is observed that the performance of RTK is unsatisfactory. Due to the prevalence of NLOS, multipath, and cycle slips in urban scenarios, RTK struggles to solve the integer ambiguity and often fails to even provide a float solution. Moreover, in cases where no solution is obtained, RTKLIB resorts to using the solution from the previous frame that had a valid solution. On the other hand, the DLS method demonstrates greater stability. However, our proposed DBCLS method outperforms both RTK and DLS, exhibiting the best performance among the three methods. In the KLT1 dataset, where DLS already performs well, our method does not surpass it significantly. However, in the KLT2 dataset, our method shows a substantial improvement over the other two methods. In the deep urban scenario of Whampoa, none of the results are satisfactory. Fig. 10

presents the predicted bias and the corresponding ground truth for each satellite in each frame across the three datasets.

The correlation coefficient  $r$  in the figure is a metric used to assess the relationship between the predicted results and the ground truth. It is calculated using the following formula:

$$r(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}} \quad (22)$$

The correlation coefficients for the three results are depicted in Fig. 10, and their values are 0.756, 0.815, and 0.634, respectively. Based on these results, it can be observed that the proposed method performs well in the KLT1 and KLT2 datasets, which have similar scenarios to the KLT3 dataset used for training. However, in the deep urban Whampoa dataset, the ground truth bias can reach up to 400 meters, while the maximum bias encountered during training on the KLT3 dataset is approximately 40 meters. As a result, the model does not provide a significant improvement in this challenging scenario.

It should be noticed that that an outlier, labeled as A, is present in Fig. 9 (a), where the DBCLS result is notably worse compared to DLS and RTK. The corresponding frame has a timestamp of 1623296321. The NLOS detection result and predicted bias for this frame are displayed in Fig. 11 and Table 3.



Fig. 11 The NLOS detection result of the outlier A.

Table 3. The detail of the predicted bias in outlier frame

Satellite	Predicted Bias(m)	Ground Truth(m)	Satellite	Predicted Bias(m)	Ground Truth(m)
G01	44.77	5.96	C07	29.69	12.74
G03	44.4	5.3	C08	-6.79	11.32
G07	41.25	3.21	C11	46.05	11.61
G14	28.93	6.85	C13	42.31	10.85
G17	41.55	8.83	C23	39.45	8.85
G22	45.96	10.44	C25	35.02	10.6
G28	41.16	6.11	C28	44.01	18.37
G30	39.85	10.05			

Fig. 11 reveals the presence of several trees next to the vehicle, with leaves obstructing the satellites. The ground truth values in Table 3 confirm that the leaves indeed have an impact on the positioning. Furthermore, the model predicted a bias significantly larger than the ground truth. This suggests a similar situation may have occurred in the KLT3 dataset, where the bias was exceptionally large for that particular frame.



## V. CONCLUSION

In this study, we introduced a fisheye camera-aided NLOS detection method. By leveraging the state-of-the-art segmentation network Swin Transformer, we achieved high accuracy, precision, and recall rates of approximately 99% for NLOS detection on images. This approach can be used independently for NLOS ground truth labeling. Additionally, we proposed a DNN architecture for pseudorange bias correction, which considers both the chosen satellite and master satellite. The network underwent evaluation on three distinct datasets, precluding data leakage, and showcased enhanced efficacy relative to traditional approaches. This underscores the notable generalization potential of the proposed technique on our dataset; however, additional datasets warrant examination. Moreover, RTKLIB's Python interface proved instrumental during our experiments.

However, it should be noted that while the NLOS detection on the image yielded accurate results, further exploration is required to determine if it corresponds to actual NLOS in the physical environment. Additionally, the current DNN architecture focuses on a single satellite in one frame, and future research should consider incorporating additional information and time sequences.

## ACKNOWLEDGMENT

We sincerely thank Dr. Guohao Zhang, Ivan, Feng Huang from Intelligent Positioning and Navigation Laboratory for the help of ground truth providing and 3DMA sky mask generating.

## REFERENCES

- [1] K. A. Redmill, S. Upadhyaya, A. Krishnamurthy, and U. Ozguner, "A lane tracking system for intelligent vehicle applications," in *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 01TH8585)*, 2001: IEEE, pp. 273-279.
- [2] T. Takasu and A. Yasuda, "Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB," in *International symposium on GPS/GNSS*, 2009, vol. 1: International Convention Center Jeju Korea Seogwipo-si, Korea, pp. 1-6.
- [3] W. Wen and L.-T. Hsu, "Towards robust GNSS positioning and real-time kinematic using factor graph optimization," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021: IEEE, pp. 5884-5890.
- [4] P. Groves, "GNSS solutions: Multipath vs. NLOS signals. how does Non-Line-of-Sight reception differ from multipath interference," *Inside GNSS Magazine*, vol. 8, no. 6, pp. 40-42, 2013.
- [5] L.-T. Hsu, "Analysis and modeling GPS NLOS effect in highly urbanized area," *GPS solutions*, vol. 22, no. 1, p. 7, 2018.
- [6] G. A. McGraw, P. D. Groves, and B. W. Ashman, "Robust positioning in the presence of multipath and NLOS GNSS signals," *Position, navigation, and timing technologies in the 21st century: integrated satellite navigation, sensor systems, and civil applications*, vol. 1, pp. 551-589, 2020.
- [7] A. Angrisano, S. Gaglione, and C. Gioia, "RAIM algorithms for aided GNSS in urban scenario," in *2012 Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS)*, 2012: IEEE, pp. 1-9.
- [8] T. Walter and P. Enge, "Weighted RAIM for precision approach," in *PROCEEDINGS OF ION GPS*, 1995, vol. 8, no. 1: Institute of Navigation, pp. 1995-2004.
- [9] V. Uaratanawong, C. Satirapod, and T. Tsujii, "Evaluation of multipath mitigation performance using signal-to-noise ratio (SNR) based signal selection methods," *Journal of Applied Geodesy*, vol. 15, no. 1, pp. 75-85, 2021.
- [10] Z. Zhang, B. Li, Y. Gao, and Y. Shen, "Real-time carrier phase multipath detection based on dual-frequency C/N0 data," *GPS Solutions*, vol. 23, pp. 1-13, 2019.
- [11] P. D. Groves, "Shadow matching: A new GNSS positioning technique for urban canyons," *The Journal of Navigation*, vol. 64, no. 3, pp. 417-430, 2011.
- [12] L.-T. Hsu, Y. Gu, and S. Kamijo, "3D building model-based pedestrian positioning method using GPS/GLONASS/QZSS and its reliability calculation," *GPS solutions*, vol. 20, pp. 413-428, 2016.
- [13] W. W. Wen, G. Zhang, and L.-T. Hsu, "GNSS NLOS exclusion based on dynamic object detection using LiDAR point cloud," *IEEE transactions on intelligent transportation systems*, vol. 22, no. 2, pp. 853-862, 2019.
- [14] W. Wen, G. Zhang, and L. T. Hsu, "Correcting NLOS by 3D LiDAR and building height to improve GNSS single point positioning," *Navigation*, vol. 66, no. 4, pp. 705-718, 2019.
- [15] J.-i. Meguro, T. Murata, J.-i. Takiguchi, Y. Amano, and T. Hashizume, "GPS multipath mitigation for urban area using omnidirectional infrared camera," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, pp. 22-30, 2009.
- [16] H. Yao, Z. Dai, W. Chen, T. Xie, and X. Zhu, "GNSS Urban Positioning with Vision-Aided NLOS Identification," *Remote Sensing*, vol. 14, no. 21, p. 5493, 2022.
- [17] T. Suzuki and N. Kubo, "N-LOS GNSS signal detection using fish-eye camera for vehicle navigation in urban environments," in *Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014)*, 2014, pp. 1897-1906.
- [18] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603-619, 2002.
- [19] X. Bai, W. Wen, and L. T. Hsu, "Using Sky-pointing fish-eye camera and LiDAR to aid GNSS single-point positioning in urban canyons," *IET Intelligent Transport Systems*, vol. 14, no. 8, pp. 908-914, 2020.
- [20] A. Vaswani *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [21] M. Chen *et al.*, "Generative pretraining from pixels," in *International conference on machine learning*, 2020: PMLR, pp. 1691-1703.
- [22] Z. Liu *et al.*, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10012-10022.
- [23] W. W. Wen and L.-T. Hsu, "3D LiDAR aided GNSS NLOS mitigation in urban canyons," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 18224-18236, 2022.
- [24] W. Wen and L.-T. Hsu, "3D LiDAR aided GNSS real-time kinematic positioning," in *Proceedings of the 34th International Technical Meeting of The Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, 2021, pp. 2212-2220.
- [25] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba, "Scene parsing through ade20k dataset," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 633-641.
- [26] U. Vollath, H. Landau, and X. Chen, "Network RTK—concept and performance," in *Proceedings of the GNSS Symposium, Wuhan, China*, 2002, pp. 27-30.
- [27] L.-T. Hsu *et al.*, "UrbanNav: An open-sourced multisensory dataset for benchmarking positioning algorithms designed for urban areas," in *Proceedings of the 34th International Technical Meeting of The Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, 2021, pp. 226-256.
- [28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [29] G. Zhang, W. Wen, B. Xu, and L.-T. Hsu, "Extending shadow matching to tightly-coupled GNSS/INS integration system," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4979-4991, 2020.