PositionNet: CNN-based GNSS Positioning in Urban Areas with Residual Maps

Penghui Xu1, Guohao Zhang1, Bo Yang2, Li-Ta Hsu*1

1 Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong, China; 2 The Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China

ABSTRACT

Multipath and non-light-of-sight (NLOS) reception greatly deteriorate the GNSS positioning accuracy in urban areas. Currently, most of the attempts using machine learning in this area focus on signal status prediction (LOS, multipath, NLOS). This paper exploits the capability of deep learning in multipath/NLOS mitigation. A new input feature, the single-differenced residual map is proposed, which has a high correlation with the user location and is very effective in multipath/NLOS mitigation. Combining the domain knowledge in GNSS, features of residual maps from different satellites are extracted by the proposed network and generate the heat map to indicate the user location. The proposed network can significantly improve the positioning accuracy of 84% of the epochs in the dense urban to 5-meter level. In addition, our network has a superior generalization ability, reducing the error of 90% of the epochs to 7 meter level in a new scenario.

Keywords: GNSS, multipath, NLOS, deep learning

1. INTRODUCTION

Global navigation satellite system (GNSS) signal can be blocked or reflected by buildings in urban areas, leading to degradation of its positioning accuracy. Such phenomena can be categorized into three cases [1]. The most straightforward is the complete signal blockage with no signal acquired and tracked. The second is the none-line-of-sight (NLOS) reception that only the reflected signal with a delay can be received. The last is the multipath effect, where both the direct and the reflected signals are received and cause interaction. Recently, machine learning has been used more and more frequently to address the multipath and NLOS problem. To generate a better pseudorange measurement, the neural network was deployed in the delay lock loop to mitigate the multipath effect [2, 3]. For signal status prediction, correlator output was used to detect NLOS signal with the support vector machine [4], convolutional neural network (CNN) [5], and multilayer perceptron (MLP) [6], respectively. And the gradient boosting decision tree (GBDT) [7], SVM [8], and CNN [9] were utilized to predict the signal status with RINEX-level features. For multipath and NLOS error prediction, multipath error in L5 is predicted using CNN with correlator output, C/N0, and elevation angle [10]. In [11], the authors applied long short-term memory (LSTM) to predict both signal status and pseudorange errors. Besides, GBDT is used to generate pseudorange error prediction, which will be further used to determine the positioning scheme [12]. For position estimation, the authors estimate the position correction based on the light-of-sight (LOS) vector and pseudorange residual using Transformer [13]. In [14], the CNN-LSTM model was utilized to deal with sequential coordinate values and generate position correction. Among all these methods, although positioning estimation attracts less attention than signal status prediction and error prediction, it is the most straightforward approach because it directly estimates or corrects the positioning level solution.

This paper proposed a neural network (PositionNet) for position estimation in a heat map. The motivation for using the heat map to indicate the user position is that the possible position should have a high correlation with the spatial pattern of residual maps. For spatial patterns, it is common to utilize CNN or a vision Transformer to serve as the feature extractor in deep learning. The main difference between these two is that CNN has stronger inductive biases: locality and translation equivariance [15]. With these two biases, CNN is easier to be trained than a vision Transformer [16]. And considering the limited volume of the training data, CNN is selected as the backbone for both encoder and decoder. Most related work in GNSS is the example in [17]. In [17], they introduced the range residual map as the CNN network input and treated the possible position as the regression problem in the local coordinate frame. Their method is combined with the 3D mapping aided (3DMA) GNSS to generate the final solution.

In our proposed method, we utilize the range residual map (Res Map), the single-differenced residual map (SDRes Map), and C/N0 as the input to generate the heat map via CNN to indicate the user position. The forward process of the neural network can

* Corresponding author: Li-Ta Hsu is with the Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University (e-mail: lt.hsu@polyu.edu.hk). Postal Address: QR828, AAE, PolyU, Hung Hom, Kowloon, Hong Kong.
be simplified into two parts: the encoding part and the decoding part. In the encoder, the input features of each satellite will go through the same feature extractor (CNN kernel) under the assumption that each satellite is equivalent. In the decoder, the extracted features from the encoder integrate together and resume to the original width and height, and eventually generate the heat map. The residual map indicates the satellite azimuth angle, elevation angle, least-squares solution, and the distance between the measurement and the expected range. Besides Res Map, the SDRes Map is developed and demonstrated to be the key to better positioning performance. SDRes Map is the combination of the Res Map of the current and master satellite (highest elevation angle one), indicating the user position. For the output, most works in deep learning regard the position estimation as the correction term regression problem. Inspired by the work in objection detection, which treats the objects in the image as points [18], this paper considers the user position as a point in a probability heat map. The experiment shows that the proposed neural network can greatly improve positioning accuracy in dense urban areas and have good generalization performance.

The contributions of this work are as below:

- Proposed the SDRes Map for deep learning, the geometry of which has a clear and straightforward relationship with the user location and multipath/NLOS phenomena.
- Developed the PositionNet, which can improve the positioning accuracy of the urban area with a remarkable result and have good generalization ability.
- Investigate the PositionNet, showing that the successful mitigation of multipath/NLOS is mainly due to the SDRes Map features from the good observation.

The rest of the paper is organized as below: Secion 2 introduces the calculation of Res Map and SDRes Map as well as their geometrical property. Secion 3 detailly introduces the network structure, label, and loss function. Secion 4 shows the experiment setup, result, and analysis. Secion 5 is the conclusion.

2. RESIDUAL MAPS

Pseudorange describes the distance between a GNSS receiver and satellite, which includes actual distance in between, receiver clock bias, satellite clock bias, and atmospheric errors. It is the most relevant measurement for positioning with trilateration. However, using this measurement directly may not be effective in machine learning because the scale and the numerical variation of the pseudorange are large.

2.1. Range Residual Map

The Res Map describes the consistency between the measured range and the expected range corresponding to each location in a certain area [17]. The range residual $\epsilon_{ijZ}$ between receiver $i$ and satellite $j$ in a sampling location $Z$ can be derived by:

$$\epsilon_{ijZ} = \rho_{ij} - \delta_i - D_{ijZ}$$

$$\hat{\rho}_i = \begin{bmatrix} \hat{x}_i \\ \hat{\delta}_i \end{bmatrix} = (G^T G)^{-1} G^T \hat{p}$$

where $\rho_{ij}$ is the measured pseudorange with atmospheric and satellite clock bias correction [19]. $\delta_i$ is the estimated receiver clock bias, $D_{ijZ}$ is the expected range calculated between a satellite and position $Z$. $\hat{\delta}_i$, as one of the unknown parameters in $\hat{p}$, is the estimated receiver clock bias by the least-squares method using Eq. (2), where $\hat{p}$ is a vector of the pseudorange measurements with atmospheric and satellite clock bias correction of different satellites. $\hat{x}$ is the estimated position. $G$ is the geometrical matrix containing the unit LOS-vector of each satellite. Assuming that the user is on the ground, the digital terrain model of Hong Kong is used to calculate the $D_{ijZ}$. Notice that the range residual actually contains information of the receiver clock bias error and multipath/NLOS error. By sampling the locations near the least-squares estimation, the range residual map can be constructed. The Res Map highly corresponds to the geometry of the satellite, as shown in Fig. 1. The grids where the residual is equal to zero can be regarded as a line. And the line (indeed a curve with super large radian) is the intersection of the spherical surface by satellite ranging measurement with the ground plane. The direction of the lines indicates the azimuth angle of the satellite. The width of the line is determined by the elevation angle, where a higher elevation angle will result in a wider line. The lines from different Res Maps in one epoch can indicate the rough user position.
2.2. Single-Differenced Residual Map

The SDRes Map is the differential of two Res Maps, aiming to exclude the receiver clock bias. The reason is that the estimated receiver clock bias in Res Map may not be accurate. The single-differenced range residual $\Delta \varepsilon_{m, i}^{m, i}$ with respect to master satellite and $i$ satellite in a location $Z$ can be derived by:

$$\Delta \varepsilon_{m, i}^{m, i} = (\tilde{\rho}_m^m - \tilde{\rho}_i^i) - (D_m^Z - D_i^Z)$$  \hspace{1cm} (3)

where $\tilde{\rho}_m^m$, $\tilde{\rho}_i^i$ is pseudorange measurement with atmospheric and satellite clock bias corrections of the master satellite (using the highest elevation angle one) and $i$ satellite, respectively. $D_m^Z$ and $D_i^Z$ is the expected range calculated between the master satellite and position $Z$, and $i$ satellite between position $Z$, respectively. Each constellation will select one master satellite. Using the highest elevation angle satellite as the master satellite is because it is less likely to be blocked by the buildings. Recall that the pseudorange can be expressed as

$$\rho_{m, i} = R_m^i + \delta_l$$  \hspace{1cm} (4)

Substituting Eq. (4) of the master satellite and current satellite $i$ to Eq. (3), we can obtain:

$$\Delta \varepsilon_{m, i}^{m, i} = (R_m^i - R_i^i) - (D_m^Z - D_i^Z)$$  \hspace{1cm} (5)

where $R_m^i$, $R_i^i$, $D_m^Z$, $D_i^Z$ is coordination of the master satellite, $i$ receiver, and $i$ satellite in the ECEF frame, respectively. $x_Z$ is the ECEF coordinate of sampling location $Z$. Similar to Res Map, the graphic pattern for location $\Delta \varepsilon_{m, i}^{m, i} = 0$ is also a line. Substituting Eq. (6), (7), (8), and (9) to Eq. (5) at the special location where $\Delta \varepsilon_{m, i}^{m, i} = 0$, Eq. (5) can be written as:

$$\|x_m^n - x_i^i\| - \|x_i^i - x_i^i\| = \|x_m^n - x_Z\| - \|x_i^i - x_Z\|$$  \hspace{1cm} (10)

For different sampling point $Z$, there is a trivial solution in Eq. (10), which is $x_Z$ equal to $x_i^i$. That means the user location is just at the lines of SDRRes Maps. And when there are multiple SDRRes Maps that do not suffer from multipath/NLOS error, the intersection of lines can indicate the user location, as shown in Fig. 2. Therefore, when the master satellite is the LOS satellite, and there are enough measurements with small pseudorange errors in one epoch, the user position is possible to be determined via the graphic pattern of the SDRRes Maps.
3. PROPOSED METHOD

Based on the residual maps, in the positioning network, we consider each satellite contains part of the information for the final positioning solution. To deal with the varying measurement of the satellites, we will firstly do the padding in the satellite’s dimension. So the input shape of one sample is \((S, C, H, W)\), where \(S\) stands for satellite number, which is 30 (if the observation is less than 30, the rest will be padded with all zeros. In our case, the maximum number of observed satellite is less than 30. If the observation is more than 30, this value can be set higher). \(C\) denotes the feature map number, in our case is four, including one Res Map of the selected satellite, one Res Map of the master satellite, one SDRes Map, and one \(C/N_0\) map. Besides the Res Map of the current satellite, it is natural to use Res Map of the master satellite together because the SDRes Map is the combination of these two. All feature maps have a shape of \(H \times W\), where \(H\) and \(W\) are both 101. In this study, candidate positions to construct Res Map and SDRes Map are sampled with 1 meter resolution in the east-west and north-south direction of the ENU frame, the center of which is the least-squares solution. In total, a map with a shape of 101 \(\times\) 101 represents the area of 101m \(\times\) 101m. The \(C/N_0\) map is filled with the value of \(C/N_0\) of the current measurement. The Res Maps and SDRes Maps are rescaled by dividing by 50, and \(C/N_0\) maps are rescaled by dividing by 45.

The detailed network architecture is similar to U-Net [20], as shown Fig. 3. The network consists of two parts: encoding network (extract features from different satellite inputs) and decoding network (retrieve the features to output size and generate a heat map). For the encoding network, assuming all the satellite measurements are equivalent, shared convolutional layers are utilized to extract features from the varying measurements. The overall encoding network is like that in Very Deep Convolutional Networks (VGG) [21], excepting that we do the same feature extraction to the features of all satellites (in the image case, they do not have dimension \(S\)). In order to enable the network to capture the large-scale spatial correlation of the residual maps, the network increases the receptive field by downsampling the feature map with the convolutional layer of stride 2 (minimum to 0.25x, where the resolution is equivalent to 4 meters/grid). The MeanPooling, as a down-sampling method, is utilized along the \(S\) axis to integrate the features of different satellites. The activation function is exponential linear units (ELU) [22]. For the decoding network, the transposed convolution is used to retrieve the feature map to output size. The high-resolution extracted features from the encoder (blue blocks after the MeanPooling) will combine with the transposed convolutional result for localization through the residual connection (blue arrow in Fig. 3). We hypothesize that the final position should have a correlation with the local pattern of the input residual maps, so the heat map of same \(H\) and \(W\) is generated to represent the user position.
The user position can be represented as a 2D tensor, where the ground truth location has the value of one, while the other location is zero. Label smoothing is made based on the horizontal dilution of precision (HDOP) and the variance of the pseudorange using Gaussian distribution in the east and north direction. The HDOP value indicates how sensitive the positioning solution is related to the pseudorange variance. The probability density function (PDF) of each grid is derived by

$$f(\mathbf{x}) = \frac{1}{2\pi \det \mathbf{\Sigma}} \exp \left( -\frac{1}{2} (\mathbf{x} - \mathbf{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \mathbf{\mu}) \right)$$

(11)

where $\mathbf{x}$ is the 2D location in the horizontal plane of the ENU frame (origin is the least-squares solution), $\mathbf{\Sigma}$ is the covariance matrix of the position error in the horizontal plane of the ENU frame. $\mathbf{\mu}$ is the ground truth location in the horizontal plane of the ENU frame. $\mathbf{\Sigma}$ is calculated using Eq. (12) and Eq. (13), where $\mathbf{G}$ is the geometrical matrix in the case with one constellation, $\sigma$ is the is pseudorange variance, which is 7.03m [19] in our case. The label is obtained by normalizing the calculated PDF. The reason for using label smoothing is that for machine learning, label smoothing can help the training process become easier [23]. In our case, label smoothing enables the network to have a lower loss while the prediction approaches the ground truth location, instead of the unified value. And this will encourage the network to find alternative solutions near the ground location. And for GNSS, even in the open-sky scenario, the position solution has the variance, which is determined by the pseudorange variance and the HDOP value. Considering the variance of position renders the network have better tolerance to the case where the satellite geometry is poor. The loss function for training is pixel-level logistic regression [18] using focal loss [24] derived by

$$\text{Loss} = \frac{-1}{N} \sum_n \left( \frac{1 - \hat{Y}_n}{\alpha} \log (\hat{Y}_n) \right) \text{ if } Y_n = 1$$

$$= \frac{-1}{N} \sum_n \left( \frac{1 - \hat{Y}_n}{\alpha} \log \left( 1 - \hat{Y}_n \right) \right) \text{ otherwise}$$

(14)

where $\hat{Y}_n$ is the predicted possibility of a grid, showing how this gird is likely to be the true location, $Y_n$ is the labeled probability of a grid. $\hat{Y}$ and $Y$ are maps of probability with the shape of $H$ and $W$, values varying from 0 to 1. $N$ is the total grid number of the predicted heat map, which is 10201 ($101 \times 101$) in our case. Subscript $n$ denotes different grids. We use $\alpha = 2$ and $\beta = 4$ in the experiment.
4. EXPERIMENT RESULTS AND ANALYSIS

4.1. Experiment Setup

The GNSS data used are collected by the low-cost receiver u-blox F9P with a standard patch antenna in Tsim Sha Tsui, Whampoa, and Mong Kok in Hong Kong. 80% of the data in each scenario are used for training and 20% for validation. Ground truth locations are collected by the Real-time kinematic INS integrated solution from an NovAtel SPANCP. The training uses Adam optimizer [25] with a learning rate of 0.0003. A 3x3 kernel with the value of one will be used to calculate the confidence level of each grid of the output as below:

\[ p_{x,y} = \sum_{row} \sum_{column} K \cdot B_{x,y} \]  

(15)

\[ K = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \]  

(16)

\[ B_{x,y} = \begin{bmatrix} \hat{Y}_{x-1,y-1} & \hat{Y}_{x-1,y} & \hat{Y}_{x-1,y+1} \\ \hat{Y}_{x,y-1} & \hat{Y}_{x,y} & \hat{Y}_{x,y+1} \\ \hat{Y}_{x+1,y-1} & \hat{Y}_{x+1,y} & \hat{Y}_{x+1,y+1} \end{bmatrix} \]  

(17)

where \( K \) is the kernel of shape 3 by 3 with value of one. \( B_{x,y} \) is part of the predicted heat map \( \hat{Y} \) with shape 3 by 3. The grid location with highest confidence level will be treated as the positioning solution from the PositionNet. And the confidence level at this special grid is used to represent the confidence level of this heat map, written as \( p_{max} \), where \( p \) is the set of all \( p_{x,y} \)

\[ p_{max} = \max(p) \]  

(18)

4.2. Validation And Testing Experiment Result

Table 1 shows the positioning results in the validation set. The result shows that the PositionNet can effectively reduce the positioning error. Given a confidence level threshold of 0.1, the error in 84% of the epochs can be improved to 5m level. Besides the positioning improvement, noticing that the least-squares positioning solution tends to be better when the \( p_{max} \) become larger, which means the \( p_{max} \) can also be used as the indicator for the overall measurement quality. For example, for the epochs where \( p_{max} \geq 0.4 \), the average positioning error of the least-squares solution is 4.3 m, which is much better than the average positioning error of all epochs (16.84 m). The validation and testing results are displayed in the map, as shown in Fig. Figure 4. Fig. Figure 4 shows that the PositionNet can produce a good positioning solution for most cases. What need to be noticed is that for some epochs in Mong Kok and Kowloon Bay, the PositionNet does not perform well. It is because the least-squares solution is too far away from the ground truth location, such that the ground truth location is out of the sampling area of the inputs (101x101 m). If the threshold is set higher than 0.1, these epochs can be easily filtered out because the \( p_{max} \) of these epochs is very small. And this is because the input residual maps drift a lot for those cases and cannot provide enough positioning information.

To evaluate the generalization ability of the model, an additional experiment is conducted in a different environment (Kowloon Bay) with 2281 epochs of data. The result is shown in Table 2. The average error of the least-squares of this scenario is larger than that in the testing dataset due to poor satellite geometry with limited satellite numbers for positioning. Even so, a large improvement can be observed even without fine-tuning the model. With the threshold of 0.1, the error in 90% of the epochs can be reduced to 7m level. Importantly, the model can also maintain good performance for epochs with good measurement qualities (like open-sky). The evidence is that when the threshold is set higher, the original positioning solution becomes better. For the cases where the least-squares solution does good, they have a very high possibility that they can be treated as open-sky cases. And for these cases, the model still has improvement.

4.3. SDRes Map Ablation Experiment Result

To investigate the effectiveness of the SDRes Map, we ablate the proposed network with different inputs. In the ablation experiment, the input containing only Res Map and \( C/N_0 \) is used. The result is shown in Table 3. Comparing Table 1 and Table 3, it is clear that the SDRes Map can greatly improve the positioning accuracy, which means the multipath/NLOS error can be significantly mitigated through the information contained in SDRes Map.
Table 1 PositionNet results in the validation set. The table shows the mean absolute positioning error in meter and improvement percentage based on different threshold values \( P \). Available Epoch represents the number of epochs and their percentage that fulfill the criteria (total 894 epochs).

<table>
<thead>
<tr>
<th>Threshold (( P ))</th>
<th>PositionNet Error (( p_{\text{max}} \geq P )) (m)</th>
<th>Available Epoch (( p_{\text{max}} \geq P )) (% of total 894)</th>
<th>Least-Square Error (( p_{\text{max}} \geq P )) (m)</th>
<th>Available Epoch (( p_{\text{max}} \leq P )) (% of total 894)</th>
<th>Least-Square Error (( p_{\text{max}} \leq P )) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7.57 (55%)</td>
<td>894 (100%)</td>
<td>16.84</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>0.1</td>
<td>5.49 (57%)</td>
<td>754 (84%)</td>
<td>12.84</td>
<td>140 (16%)</td>
<td>38.40</td>
</tr>
<tr>
<td>0.2</td>
<td>3.49 (69%)</td>
<td>472 (53%)</td>
<td>11.13</td>
<td>422 (47%)</td>
<td>23.23</td>
</tr>
<tr>
<td>0.3</td>
<td>2.82 (70%)</td>
<td>241 (26%)</td>
<td>9.54</td>
<td>653 (73%)</td>
<td>19.54</td>
</tr>
<tr>
<td>0.4</td>
<td>2.56 (38%)</td>
<td>54 (6%)</td>
<td>4.30</td>
<td>840 (94%)</td>
<td>17.65</td>
</tr>
</tbody>
</table>

Figure 4 Validation and testing result displayed in the map.

Table 2 PositionNet result in the testing set. The table shows the mean absolute positioning error in meter and improvement percentage based on different threshold values \( P \). Available Epoch represents the number of epochs and their percentage that fulfill the criteria (total 2281 epochs).

<table>
<thead>
<tr>
<th>Threshold (( P ))</th>
<th>PositionNet Error (( p_{\text{max}} \geq P )) (m)</th>
<th>Available Epoch (( p_{\text{max}} \geq P )) (% of total 2281)</th>
<th>Least-Square Error (( p_{\text{max}} \geq P )) (m)</th>
<th>Available Epoch (( p_{\text{max}} \leq P )) (% of total 2281)</th>
<th>Least-Square Error (( p_{\text{max}} \leq P )) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11.39 (41%)</td>
<td>2281 (100%)</td>
<td>19.25</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>0.1</td>
<td>7.77 (47%)</td>
<td>2046 (90%)</td>
<td>14.70</td>
<td>235 (10%)</td>
<td>58.90</td>
</tr>
<tr>
<td>0.2</td>
<td>5.92 (50%)</td>
<td>1433 (63%)</td>
<td>11.94</td>
<td>848 (37%)</td>
<td>31.61</td>
</tr>
<tr>
<td>0.3</td>
<td>4.84 (34%)</td>
<td>792 (35%)</td>
<td>7.30</td>
<td>1489 (65%)</td>
<td>25.61</td>
</tr>
<tr>
<td>0.4</td>
<td>4.82 (15%)</td>
<td>241 (11%)</td>
<td>5.66</td>
<td>2040 (89%)</td>
<td>20.86</td>
</tr>
</tbody>
</table>
Table 3 Neural network positioning results in using only Res Map and C/N₀ in the validation set.

<table>
<thead>
<tr>
<th>Threshold (P)</th>
<th>PositionNet Error (ₚ₉₉₉ ≥ P) (m)</th>
<th>Available Epoch (ₚ₉₉₉ ≤ P)</th>
<th>Least-Square Error (ₚ₉₉₉ ≥ P) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12.80 (24%)</td>
<td>894 (100%)</td>
<td>16.84</td>
</tr>
<tr>
<td>0.1</td>
<td>4.34 (20%)</td>
<td>334 (37%)</td>
<td>5.41</td>
</tr>
<tr>
<td>0.2</td>
<td>2.19 (1%)</td>
<td>9 (1%)</td>
<td>2.19</td>
</tr>
<tr>
<td>0.3</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>0.4</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

4.4. Analysis

From the experiment results, it is clear that the network can mitigate the multipath/NLOS. To further study how the network capture the correlation between different satellite features and the output heat map, the importance of different inputs corresponding to the current output is investigated. The loss calculation in Eq. (14) can be expressed as the non-linear function

\[
\text{Loss} = f(\hat{Y}, Y)
\]

(19)

where \(\hat{Y}\) is the prediction, \(Y\) is the label, \(I\) is the input, \(w_0\) is the trained network parameter. Given an input \(I_0\) and its corresponding label \(Y_0\), the gradient of the output at \(I_0\) and \(Y_0\) can be expressed as

\[
\frac{\partial \text{Loss}}{\partial I}
\]

\(I_0, Y_0\)

which can be calculated using the first-order Taylor expansion. Where \(w\) is the partial derivative of the loss with respect to input \(I\) at \(I_0\) (a specific input) given the label \(Y_0\), \(L_{o}s\) is calculated using Eq. (14). \(w\) is the tensor have the same shape as the input \(I_0\). The derivative of loss with respect to the input is called saliency map [26]. The larger value of \(w\) means the loss is more sensitive to the corresponding value in \(I_0\). In other words, larger \(w\) means the corresponding input show higher significance. Noticing that the \(\text{Loss}\) can also be expressed as the summation form as

\[
\text{Loss} = \sum_n \text{Loss}_n = \sum_n f(\hat{Y}_n, Y_n)
\]

(22)

where subscript \(n\) denotes different grids of the heat map. And the grid at prediction \(\hat{Y}_n\) with the value \(p_{\text{max}}\) is what we care most, because it indicates the most possible user location for the given input. So instead of investigating \(w\) of the whole heat map, we calculate the partial derivative of loss at the grid with the highest confidence level \(p_{\text{max}}\). The derivative is written as

\[
\frac{\partial \text{Loss}_h}{\partial I}
\]

\(I_0, Y_0\)

The subscript \(h\) denotes the index where \(\hat{Y}_h = p_{\text{max}}\). The saliency map calculated using Eq. (23) also has the same shape as the input, which is \((S, C, H, W)\) for one sample. Along axis \(C\), the mean value was taken, so the saliency map becomes \((S, H, W)\). And then, the \(W\) and \(H\) are compressed by summation. So the saliency map weighting with shape \((S)\) representing the importance of each satellite is obtained. Below, three different cases will be investigated using the saliency map weighting and saliency map.

Three cases are analyzed. Case 1 is for the poor least-squares solution, while PositionNet performs well. Case 2 is that both least-squares and neural network solutions perform well. Case 3 is neither performance of least-squares nor the proposed method satisfactory. Fig. 5 shows the PositionNet output using OpenStreetMap and heat map in Case 1. Table 4 shows the saliency map weighting of some satellites in this case. The pseudorange error is calculated based on the double-differenced method using Equation (17) in [27]. Higher weighting means this satellite is more influential to the prediction location. Comparing the weighting, we noticed that the satellite with a small pseudorange error tends to have a higher weighting. And satellites with large pseudorange and master satellites tend to have less influence on the calculation. It is understandable that the master satellite input has low weight because the information of the master satellite is already encoded into other input satellites by the SDRes Map and the master Res Map. From Table 4, two interesting discoveries are found: 1) the network implicitly learned the correlation between the pseudorange error and the input. 2) the network learns to de-weight the measurement with a large pseudorange error. These two findings reveal what the network has learned and why it works to some degree. To further investigate how the relationship between bad quality measurement and weighting is built, the saliency map is plotted separately, as shown in Fig. 6. We have noticed that the SDRes Map contains the majority of the gradient of the saliency map, which means the network highly relies on the SDRes Map features, and this is consistent with the result of the ablation experiment in Table 3.
Fig. Figure 7 shows the prediction result of Case 2. In Table 5, the different weightings can also be observed in Case 2. And because the SDRes Map of each satellite (except the master ones) can provide adequate and accurate positioning information in a different direction, the varied weighting is less significant in this case. However, in Fig. Figure 8, it is noticeable that the network still overlooks the Res Map even if it is good enough. This can be because, for most of the cases in training, compared to SDRes Map, the Res Map feature is less reliable. And eventually, the network tends to focus on the SDRes Map more.

The result of Case 3 is shown in Fig. Figure 9. The error in the North-South direction is far more significant than that in the West-East one. The error is considered to occur due to the below two reasons: 1) the master satellite itself is suffering from a multipath/NLOS problem 2) the poor geometry in SDRes Map. For the first reason, from Eq. (10), it could be seen that the local pattern of the user location can be obvious only in the cases when the master satellite is a LOS satellite. For the second reason, it is clear that only two SDRes Maps (PRN 21 and 30) can provide positioning information in the north-south direction, as shown in Fig. Figure 10, and they have been distributed with large weighting, as in Table 6. However, these two SDRes Maps fail to provide accurate and redundant features. And the lack of information from other satellites further causes the network can only rely more on the master satellite, which actually does not contain enough information (the SDRes Map of the master satellite is all zero values). In Case 3, all of these issues combined together render the network failing to produce a satisfying result. To clearly see the correlation between the saliency map weighting of different satellites, pseudorange error, and output positioning error, Fig. Figure 11 is plotted. Fig. Figure 11 shows that the network successfully keeps the satellites with large pseudorange errors with relatively low weighting, and the epochs with large pseudorange errors tend to have larger output positioning errors.

Case 1

![Figure 5 Positioning solution in Case 1. Location: Whampoa. Least-squares error east and north: 30.8 m and 29.1 m. (a) Solution displayed on a map. (b) Output heat map of the PositionNet, the center of which is the least-squares solution.](image)

Table 4 Saliency map weighting of each observation for Case 1. Pr_Err is short for pseudorange error, calculated using Equation (17) in [27]. * marker after PRN means the satellite is selected as the master satellite in the current constellation. Blue, yellow, green, and orange represent GPS, GLONASS, GALILEO, and BEIDOU, respectively (same for Table 5 and Table 6).

<table>
<thead>
<tr>
<th>PRN</th>
<th>2</th>
<th>6</th>
<th>14</th>
<th>17</th>
<th>19</th>
<th>28*</th>
<th>30</th>
<th>46</th>
<th>47*</th>
<th>49</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr_Err</td>
<td>108.41</td>
<td>1.32</td>
<td>0.78</td>
<td>2.28</td>
<td>1.82</td>
<td>0</td>
<td>105.54</td>
<td>31.63</td>
<td>0</td>
<td>-1.17</td>
<td>-2.46</td>
</tr>
<tr>
<td>Saliency Map Weighting</td>
<td>0.9</td>
<td>19.27</td>
<td>10.09</td>
<td>11.41</td>
<td>12.43</td>
<td>0.94</td>
<td>0.08</td>
<td>1.72</td>
<td>1.72</td>
<td>16.6</td>
<td>12.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PRN</th>
<th>57</th>
<th>63</th>
<th>77*</th>
<th>82</th>
<th>93</th>
<th>94</th>
<th>96*</th>
<th>97</th>
<th>99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr_Err</td>
<td>1.47</td>
<td>1.46</td>
<td>0</td>
<td>1.14</td>
<td>-0.95</td>
<td>-1.81</td>
<td>0</td>
<td>-1.49</td>
<td>-1.21</td>
</tr>
<tr>
<td>Saliency Map Weighting</td>
<td>11.6</td>
<td>4.88</td>
<td>0.55</td>
<td>8.97</td>
<td>15</td>
<td>13.21</td>
<td>0.31</td>
<td>17.78</td>
<td>14.52</td>
</tr>
</tbody>
</table>
Figure 6 Saliency map of Case 1. Row 1: PRN. Row 2, 4, 6, 8: Res Map of current satellite, Res Map of master satellite, SDRes Map, $C/N_0$ Map. Row 3, 5, 7, 9: the saliency maps correspond to row 2, 4, 6, 8.
Case 2

Figure 7 Positioning solution in Case 2. Location: Tsim Sha Tsui. Least-squares error east and north: 0.8 m and 2.7 m. (a) Solution displayed on a map. (b) Output heat map of the PositionNet, the center of which is the least-squares solution.

Table 5 Saliency map weighting of each observation for Case 2.

<table>
<thead>
<tr>
<th>PRN</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>21*</th>
<th>30</th>
<th>48</th>
<th>51*</th>
<th>52</th>
<th>57</th>
<th>69</th>
<th>77</th>
<th>82*</th>
<th>93*</th>
<th>95</th>
<th>96</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr_Err</td>
<td>-1.2</td>
<td>-1.2</td>
<td>-1.9</td>
<td>0</td>
<td>-3.1</td>
<td>3.83</td>
<td>0</td>
<td>-0.8</td>
<td>-1.7</td>
<td>-1.3</td>
<td>1.9</td>
<td>0</td>
<td>0</td>
<td>4.1</td>
<td>3.3</td>
</tr>
<tr>
<td>Saliency Map Weighting</td>
<td>7.2</td>
<td>6.7</td>
<td>25.4</td>
<td>1</td>
<td>5.1</td>
<td>8.1</td>
<td>3.1</td>
<td>4.1</td>
<td>7.3</td>
<td>13.1</td>
<td>4.5</td>
<td>2.1</td>
<td>2.9</td>
<td>25</td>
<td>28.1</td>
</tr>
</tbody>
</table>

Figure 8 Saliency map of Case 2. Row 1: PRN. Row 2, 4, 6, 8: Res Map of current satellite, Res Map of master satellite, SDRes Map, $C/N_0$ Map. Row 3, 5, 7, 9: the saliency maps correspond to row 2, 4, 6, 8.
Case 3

Figure 9 Positioning solution in Case 3. Location: Mong Kok. Least-squares error east and north: -22 m and 0.1 m. (a) Solution displayed on a map. (b) Output heat map of the PositionNet, the center of which is the least-squares solution.

Table 6 Saliency map weighting of each observation for Case 3.

<table>
<thead>
<tr>
<th>PRN</th>
<th>1*</th>
<th>7</th>
<th>8</th>
<th>17</th>
<th>21</th>
<th>30</th>
<th>42*</th>
<th>53</th>
<th>71*</th>
<th>86</th>
<th>93</th>
<th>109</th>
<th>113*</th>
<th>114</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr_Err</td>
<td>0</td>
<td>1.19</td>
<td>68.61</td>
<td>-0.95</td>
<td>4.3</td>
<td>16.7</td>
<td>0</td>
<td>-2.44</td>
<td>0.77</td>
<td>0.91</td>
<td>41.16</td>
<td>0</td>
<td>9.64</td>
<td></td>
</tr>
<tr>
<td>Saliency Map Weighting</td>
<td>15.3</td>
<td>5.5</td>
<td>2.8</td>
<td>6.5</td>
<td>23.4</td>
<td>16.4</td>
<td>17.6</td>
<td>4.7</td>
<td>14.3</td>
<td>11.1</td>
<td>7.8</td>
<td>3.3</td>
<td>13.9</td>
<td>17.3</td>
</tr>
</tbody>
</table>

Figure 10 Saliency map of Case 3. Row 1: PRN. Row 2, 4, 6, 8: Res Map of current satellite, Res Map of master satellite, SDRes Map, C/A Map. Row 3, 5, 7, 9: the saliency maps correspond to row 2, 4, 6, 8.
5. CONCLUSION AND FUTURE WORK

This paper introduced the PositionNet for positioning and it yields remarkable improvement. The generalization ability of PositionNet is good even in a new scenario without fine-tuning the model. In the validation set, 84% of the epochs can be improved to 5 m level, and in the testing set, the error of 90% of the epochs can be reduced to 7 m level. In addition to improving the positioning accuracy, the PositionNet can also be used as the indicator for overall measurement quality. When the maximum confidence level of the heat map is high, the least-squares solution tends to have a small error and vice versa. Besides, the SDRes Map is proposed and demonstrated to be very effective in multipath/NLOS mitigation according to the ablation experiment. Through the investigation of some examples using the saliency map, the PositionNet is found to be able to de-weight the poor measurement based on the SDRes map. Knowing that the SDRes Map is the key feature of PositionNet, in future work, it is worth exploring how to adaptively select the master satellite in the neural network to obtain better SDRes Maps instead of just using the elevation angle.

REFERENCE


**Penghui Xu** received a B.S. degree from South China Agricultural University in 2015. In 2017, he obtained his MSc degree in mechanical engineering from The Hong Kong Polytechnic University. After that, he mainly works in machine learning algorithm development. Currently, he is pursuing a Ph.D. degree at The Hong Kong Polytechnic University. His research interests include machine learning, GNSS urban localization, and multi-sensor integration for positioning.
Guohao Zhang (Student Member, IEEE) received his bachelor’s degree in mechanical engineering and automation from University of Science and Technology Beijing, China, in 2015. He received his M.S. degree in Mechanical Engineering and his Ph.D. degree in Aeronautical and Aviation Engineering from the Hong Kong Polytechnic University. He is currently a Postdoctoral Research Fellow with the Department of Aeronautical and Aviation Engineering, the Hong Kong Polytechnic University. His research interests include GNSS urban positioning, collaborative positioning, and multi-sensor integrated navigation.

Dr. Bo Yang is an Assistant Professor in the Department of Computing at The Hong Kong Polytechnic University. He obtained his DPhil degree (2020) from the University of Oxford. His research interests lie in machine learning, computer vision and robotics.

Dr. Li-Ta Hsu (Member, IEEE) is an associate professor at Department of Aeronautical and Aviation Engineering of Hong Kong Polytechnic University. He is Limin Endowed Young Scholar in Aerospace Navigation. He received the B.S. and Ph.D. degrees in Aeronautics and Astronautics from National Cheng Kung University, Taiwan, in 2007 and 2013, respectively. He was a Visiting Researcher with the Faculty of Engineering, University College London and Tokyo University of Marine Science and Technology, in 2012 and 2013, respectively. In 2013, he won a Student Paper Award and two Best Presentation Awards from the Institute of Navigation (ION). He was selected as a Japan Society for the Promotion of Sciences Postdoctoral Fellow with the Institute of Industrial Science, The University of Tokyo and worked from 2014 to 2016. He is an Associate Fellow in the Royal Institute of Navigation. Dr. Hsu currently is members of ION and IEEE and serves as a member of the editorial board and reviewer in professional journals related to GNSS.