2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

# Looking Beneath More: A Sequence-based Localizing Ground Penetrating Radar Framework

Pengyu Zhang[1], Shuaifeng Zhi[1], Yuelin Yuan[2], Beizhen Bi[1], Qin Xin[1], Xiaotao Huang[1], Liang Shen[1]

***Abstract*— Localizing ground penetrating radar (LGPR) has been proven to be a promising technology for robot localization in various dynamic environments. However, the extreme scarcity of underground features introduces false candidate matches and brings unique challenges to this task. In this paper, we propose a sequence-based framework for LGPR to address the aforementioned issues. Specifically, we first introduce a trainable strategy to extract robust underground features in multi-weather conditions. By further using sequential information, our LGPR system can observe richer underground scene contexts, and the associated multi-frame scans could also improve the performance of underground place recognition. We demonstrate the superiority of our proposed method by comparing it against several recent state-of-the-art baseline methods applied to GPR image tasks. Experimental results on large public and self-collected datasets show that our proposed framework significantly improves the performance of various baselines in different scenarios.**

## I. INTRODUCTION

Ground penetrating radar (GPR) has been widely used in various engineering applications due to its non-contact, non-destructive advantages, as noted in prior studies [1]–[3]. Recently, robot localization by GPR has attracted extensive and increasing attention [4], [5]. Frequent changes in personnel and machine positions within indoor factories, along with seasonal and lighting variations in outdoor environments, often render traditional localization techniques ineffective. Nevertheless, GPR effectively addresses the challenge posed by dynamic surface conditions by detecting stable underground features.

Localizing ground penetrating radar (LGPR) was introduced by Cornick et al. [4], [6], [7], which compares the current GPR scan with a global database for positioning. However, it necessitates looking for the best position in a prior region of around 5m.

Evangelos et al. proposed the use of GPR A-scans to localize a GPR antenna within a known subsurface map [8]. Baikovitz et al. expanded upon this idea and employed GPR for underground mapping in unexplored areas [5]. In this research, by combining a revisit strategy with factor graph inference, they utilized GPR to calibrate IMU offset, enhancing the accuracy of motion trajectory estimation. The above localization methods based on single-channel GPR rely on significant underground features (pipelines, targets) and require the integration of other sensor information to achieve localization.

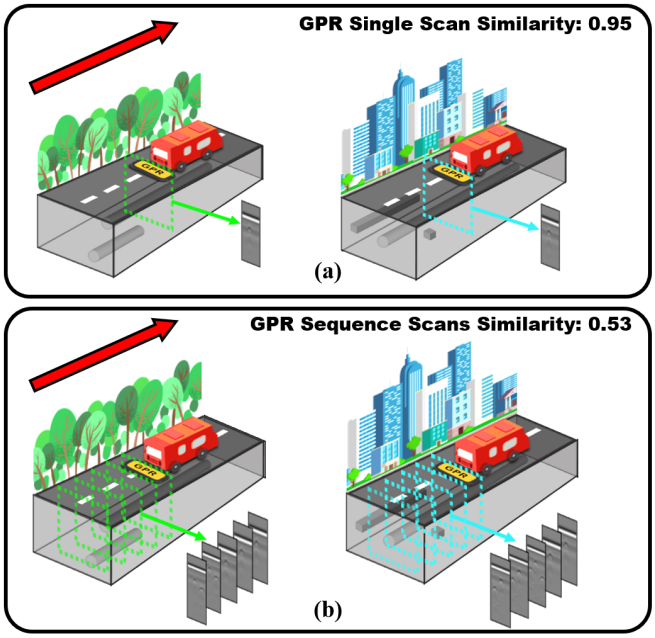


Fig. 1. LGPR system in different roads (left: Jungle Street. right: City Street). (a) Single frame GPR scan has high similarity in two completely different scenes, making it prone to incorrect matching between the two scenes. (b) Two different scenes can be effectively distinguished by looking at more underground scenes through multi-frame GPR scans.

Differential weather conditions can lead to fluctuations in underground water content, which may potentially affect the precision of LGPR measurements. [7], [9]. Some end-to-end CNN-based networks have recently been proposed to address the multi-weather challenges [9]–[11]. While these tasks use end-to-end networks for direct robot position regression, several critical issues remain inadequately addressed. These methods lack interpretability, making it challenging to identify existing deficiencies within the LGPR. Furthermore, the experimental results are based on average outcomes from specific datasets, which can pose challenges for future researchers attempting to replicate them. Therefore, these solutions are not well-suited for analyzing potential problems.

Overall, there are two main issues in LGPR. Firstly, the current LGPR research lacks a public benchmark for assessing its positioning performance in diverse scenarios. Secondly, existing technical solutions often depend on an approximate prior position or the integration of other sensor data, which poses challenges in addressing the problem of

[1]Pengyu Zhang is with the College of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan 410073, China, e-mail: zhangpengyu970524@163.com
[1]Liang Shen is the corresponding author, e-mail: shenliang16@nudt.edu.cn
[2]Yuelin Yuan works at Hikauto, Hangzhou, Zhejiang 310000, China

robot kidnapping.

In this paper, we propose a sequence matching-based framework to address the limitations of LGPR discussed earlier. Inspired by GPR's applications in various engineering fields [12]–[14]. Unlike existing methods that rely on limited GPR observations [9], [10], our method focuses on increasing LGPR system observations through sequence information. Additionally, we have designed a trainable network for extracting robust underground features in multi-weather conditions. Following is a summary of our paper's contributions:

1) We propose a sequence matching-based LGPR framework to reduce candidate matching errors effectively. Besides, we consider sequence continuity and develop a reranking mechanism for LGPR.

2) We have developed a multi-weather scan (MWS) descriptor using deep learning that exhibits robustness across different weather conditions. The proposed method outperforms a variety of competing descriptors in GPR applications.

3) The experimental results analyze the performance of our localization approach in various stages and challenging scenarios using both the GROUNDED dataset and a self-collected dataset. The code will be released to the community upon acceptance at https://github.com/Pyxel970524/SeqLGPR

## II. Related Work

LGPR achieves localization by matching the current scan with scans in the database, a principle similar to visual place recognition (VPR). In this section, we introduce the current development status of VPR and outline the potential for sequence matching in LGPR.

### A. Visual Place Recognition (VPR)

Although LGPR is still in relatively early stages of development, there have been numerous attempts to develop VPR, which can offer valuable insights.

In recent years, researchers have primarily enhanced VPR performance through two main approaches:

1) Designing global features, as seen in classic works like Netvlad, Patch-Netvlad, ETR, and others [15]–[18]. Robust global features are essential for distinguishing the current position from other locations on the global map.

2) Leveraging sequential information to enhance performance, as demonstrated by methods like SeqSLAM, SeqMatchNet, SeqVlad, and others [19]–[22], which take the continuity of robot motion into account.

The success of these methods can be credited to the wealth of information available in optical images. However, unlike visual sensors, the limited observations of GPR contain only a small number of features, making it challenging for existing metric learning methods to achieve convergence.

### B. Sequence matching in LGPR

Consider a vehicle-mounted GPR system passing through two different scenes, as illustrated in Fig. 1a. These two scenes, 'jungle' and 'urban streets', represent distinct places. In both scenes, an artificial pipeline is detected by the GPR. Interestingly, a high level of similarity exists between the GPR single scans of these two different scenes. This high similarity can cause these two distinct locations to be incorrectly identified as the same scene. The characteristic of vertical detection allows GPR to only observe a limited area. Typically, a single GPR detection can cover a real-world width of about 1m and a detection depth of approximately 3m [6], [23]. As the database expands, these factors will lead to an increasing incidence of false candidate matches in LGPR.

To tackle these challenges, an intuitive approach is to combine early GPR scans. As shown in Fig. 1b, aggregating previous multi-frame scans enables the system to capture a broader range of underground scenes. This diminishes the similarity between GPR scans from distinct locations, thereby lowering the risk of incorrect matching.

It's worth noting that, unlike visual sensors, GPR detects minimal overlapping or redundant information in adjacent frames. Each additional frame provides the robot with new information.

## III. Methodology

In this section, we introduce the process of the proposed sequence-based framework, which comprises four main components: learning-based feature extraction, construction of a similarity matrix, sequence matching and reranking. Fig. 2 illustrates the framework's pipeline.

### A. Learned *M*ulti-*W*eather *S*can (MWS) Descriptor

While underground scenes are generally more stable than above-ground scenes, variations in surface and soil moisture content due to different weather conditions can impact the collected GPR data. In this section, we introduce MWS-Net, a learning-based method designed to transform GPR scans under different weather conditions into a shared feature space.

Compared to VPR, LGPR is more sensitive to labels, making it difficult to apply VPR methodologies to LGPR. Inspired by feature transfer [24], we first set up a multi-weather template matching task that is not sensitive to labels. And migrate the network as a feature extractor for multi-weather GPR scans. The network architecture, as depicted in Fig. 4, follows a Siamese network structure with a U-shaped feature extraction network. The fusion of multi-scale downsampling and upsampling features of U-Net has been proven to be beneficial for extracting robust information of high and low frequencies in GPR images [25]–[28].

The features extracted by the two U-Net are then combined by a correlation layer. The network's final output is a score map, of which the point with the highest score represents the best matching position between the two input images. In practice, the network will calculate the similarity of all sliding sub-windows in a single evaluation.

We employ a discriminative approach, training the network on positive and negative pairs and adopting logistic loss as

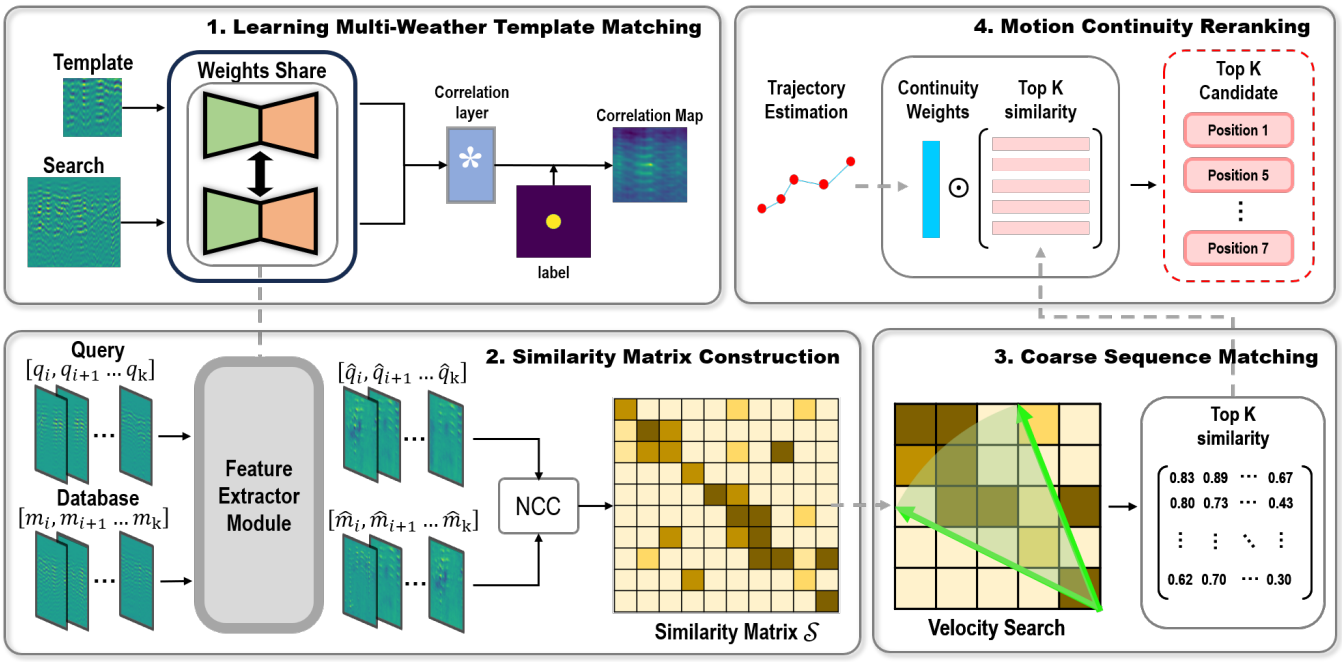


Fig. 2. **Overview of the sequence matching based LGPR framework.** The proposed framework has four components: **1**. The U-Net based Siamese network extracts multi-scale features from GPR template matching pairs to enhance the correlation map (Section III-A). **2**. The raw GPR scans are processed by **M**ulti-**W**eather **S**can (MWS) module and normalized cross-correlation (NCC) is used to match the query GPR scans with the GPR scans in the database, obtaining the similarity matrix $\mathcal{S}$ (Section III-B). **3**. For each sequence to be matched, the velocity range is limited to achieve coarse sequence matching (Section III-C).**4**. The current position is estimated by trajectory prediction and motion continuity weights are used to rerank the Top $K$ candidates. The Top 1 of reranking matches is the final prediction (Section III-D).

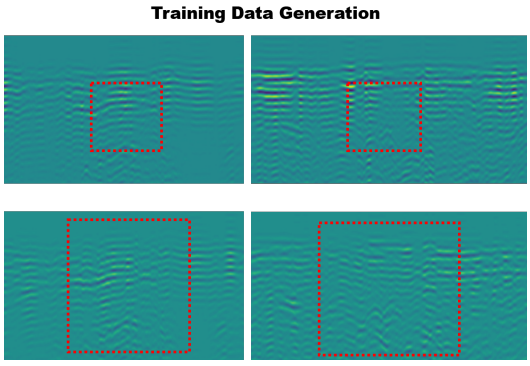


Fig. 3. Template matching training image pairs extracted from the different weather in the same route (up: template image, down: search image)

shown below:

$$l\left(g,p\right)=\log\left(1+e^{-gp}\right), \tag{1}$$

where $g$ is the ground truth binary map, and $p$ is the network prediction result. The loss of the predicted correlation map is defined as the average loss between each pixel in response map $\mathcal{M}$:

$$\mathcal{L}\left(g,p\right)=\frac{1}{|\mathcal{M}|}\sum_{(x,y)\in\mathcal{M}} l\left(g\left[x,y\right],p\left[x,y\right]\right). \tag{2}$$

Extracting image pairs from GPR datasets collected under various weather conditions. These pairs are composed of templates and search images centered around specific regions, as shown in Fig. 3. It should be noted that the training image does not need to contain specific underground targets or significant features, because our goal is to focus network training on multi-weather transformation, rather than specific target classification or detection.

### *B. Similarity Matrix Construction*

A similarity matrix is established to store the similarity scores between the current GPR query scan and the database. Various common similarity measures [3] were experimentally investigated. In on our experiments, the similarity $\mathcal{S}$ between the current query scan $q$ and each database scan $m$ was calculated using normalized cross-correlation (NCC):

$$\mathcal{S}(i,j)=\frac{\sum_{i,j}(q(i,j)-\overline{q})(m(i,j)-\overline{m})}{\sqrt{\sum_{i,j}(q(i,j)-\overline{q})^2\sum_{i,j}(m(i,j)-\overline{m})^2}}. \tag{3}$$

Each pixel coordinates $(i,j)$ in the matrix represents the similarity between the $i$-th frame of the current data and the $j$-th frame in the database. As shown in Fig.5, The closer the pixel color is to red, the higher the similarity between the corresponding two scans. Ideally, the similarity matrix should resemble a unit matrix, indicating a one-to-one correspondence between the current scan data and the scans in the database.

However, the similarity matrix generated by raw GPR scans actually displays a darker color distribution. As shown in Fig. 5 b, indicating that the query GPR scan corresponds to a sizable number of candidate matches in the map database.

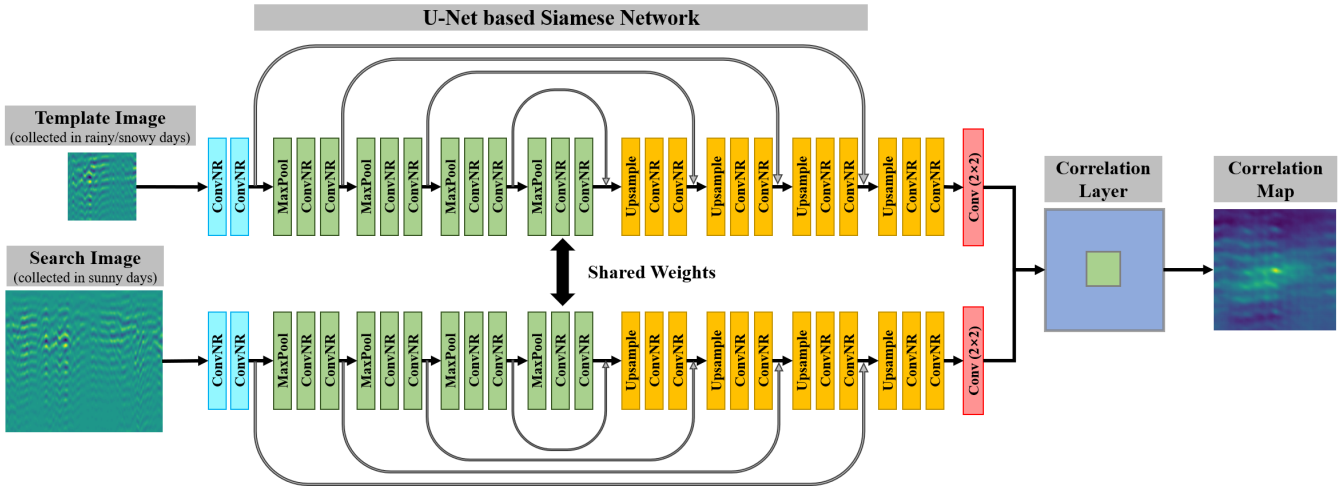


Fig. 4. Network architecture of the proposed MWS-Net. The template and search image extract multi-scale features through U-Net with the same weight, and a score map is output through a cross-correlation layer. The more obvious the center kurtosis of the score map, the better the template matching performance.

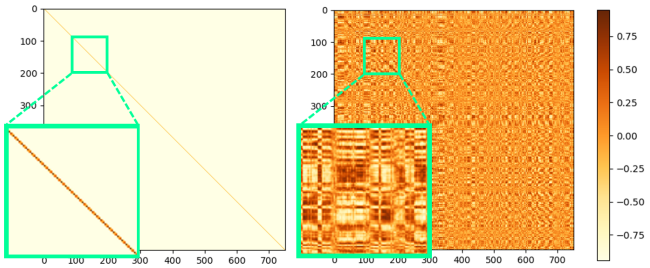


Fig. 5. (a) Ideal similarity matrix. (b) Raw LGPR scan similarity matrix.

### C. Sequence Matching

The localization in sequence-based matching is determined by comparing the correlation accumulation score, denoted as $\mathcal{C}$, between multi-frame queries and maps. A higher $\mathcal{C}$ score indicates a stronger match between these two sequences. However, it's important to note that the relationship between the query and the map may not always align with the chronological order of data collection, which can be attributed to variations in vehicle speeds during different runs. Following the notations from prior research [19], this involves defining a specific speed ratio range for the search.

Matching two sequences is represented by a trajectory line in the similarity matrix. The correlation accumulation score $\mathcal{C}$ is computed for each trajectory line by considering the similarity values it traverses from time $T-dl$ to the current time $T$:

$$\mathcal{C} = \sum_{q=T-dl}^{T} \mathcal{S}(q, d), \tag{4}$$

where $q$ corresponds to the index of the sequence data query currently collected, $dl$ is the select sequence length, and $m$ is the index in the GPR map database. The index of the map database, denoted as $d$, is defined as follows:

$$d = M + V\,(t + dl - T)\,, \tag{5}$$

where $V$ represents the trajectory velocity, which varies between $V_{min}$ and $V_{max}$ in increments of $V_{step}$. $M$ denotes the database scan index from which the trajectory originates.

Fig. 6 illustrates trajectory lines for different speeds originating from a single scan, with the green line representing the speed that yields the highest similarity summation.

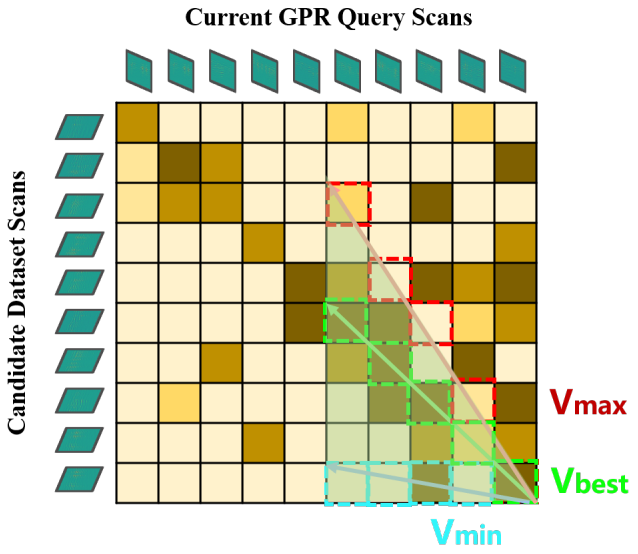


Fig. 6. Searching for best matching sequences within the space of recent scan similarity matrix.

### D. Motion Continuity Reranking

Reranking has received much attention in VPR in recent years [29], [30]. However, in GPR data, the observed scenes

are constrained, exhibiting sparse characteristics, which renders it challenging to refine local features based on global features akin to VPR. Inspired by the widespread consideration of geometric constraints in feature matching [31]. This article proposes a reranking mechanism based on motion continuity constraints, derived from sequence continuity.

Top $k$ candidate matches are coarse screened using velocity search. These candidates are subsequently reranked using continuity weights, which are determined using the positions of candidate matches ($p_i$) and the predicted current position ($p_x$). This ensures trajectory smoothness and continuity. The weights are calculated as follows:

$$w_i = e^{-\|p_i - p_x\|/\sigma}, i = 1, 2 \ldots, k, \tag{6}$$

where $\sigma$ represents a scale factor, $k$ is the Top $k$ results participating in reranking.

The current position $p_x$ can be estimated using time series prediction. In this paper, we treat the motion of GPR as moving at a uniform speed over a limited number of sequence frames, which can be represented using a simple linear regression to describe past positions [32].

$$\triangle p_x = \frac{1}{k} \sum_{i=1}^{k} \triangle p_{x-i+1}, \tag{7}$$

$$p_x = p_{x-i+1} + \triangle p_x, \tag{8}$$

where $\triangle p_x$ is the predicted displacement change. The similarity of candidate matches will be reranked by the updated correlation score $\hat{\mathcal{C}}_i$.

$$\hat{\mathcal{C}}_i = w_i * \mathcal{C}_i, i = 1, 2, ..., k. \tag{9}$$

## IV. Experiment and Evaluation

### A. Datasets

In this section, we validate our algorithm on public datasets and self datasets each of which consisted of at least three traversals of the same route but at different times.

**GROUNDED Dataset** is currently the only large public dataset for LGPR based on an 11-channel GPR [17]. The dataset includes 7 different scene roads (highway, downtown, and city); three weather conditions (sunny, rainy, and snowy); multi-lane mapping challenge (Left, Right, Center, Changing). A total of 108 runs of data.

**Self Dataset** was collected under different surface roads and conditions to enrich our evaluation, as shown in Fig. 7. To collect the data, the GPR system was mounted behind a motorcycle. The antenna array of the GPR system consisted of two-row uniform arrays with 20 radar channels (DXG 1820, 3D-RADAR). The frequency range of the GPR system is 50–3050 MHz. Self-datasets included three different types of roads and two weather conditions (sunny, and rainy)

### B. Experiment Details

- **Training**: In our experiment, Route 4 from GROUNDED[1] and Route 2 from self datasets were selected as training data. We synchronize RTK-GPS and GPR data based on timestamps and align data from different runs. The model is trained on 1 RTK-3060 GPU with Adam optimizer [33].
- **Evaluation**: 7 distinct map/query run pairs from GROUNDED[2] and self datasets (Route 1,3) are used to evaluate different methods, none of which were used for training.
- **Evaluation Metric**: In this paper, we use Recall-Rate@K as an evaluation metric, which is a widely adopted indicator in visual place recognition, as seen in previous works [15]–[17].

The parameters involved in the algorithm and experiments are summarized in Table I.

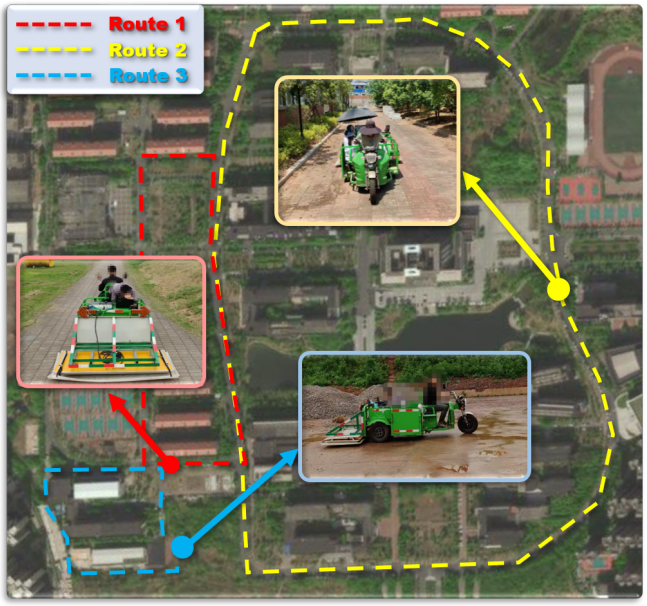


Fig. 7. Aerial photos and GPR data collection paths.

TABLE I
ALGORITHM AND EXPERIMENTAL PARAMETER SETTINGS

| Parameter | Description and Value |
|---|---|
| Train Pairs | Number of training samples. 4000 |
| Learning Rate | Neural Network Learning Rate. 1e-5 |
| Label Radius | The radius of template matching labels. 15 |
| $R_x$,$R_y$ | Resize GPR scans by bilinear interpolation. [$R_x$,$R_y$] = [400,400] |
| Vmin,Vmax | Min/Max speed ratio. Vmin=0.6, Vmax = 1.4 |
| $\sigma$ | Scale factor for reranking. $\sigma$ = 20 |
| $k$ | Top $k$ candidate matches, $k$ = 15 |

### C. Comparison of different features in LGPR

To our knowledge, there are no feature descriptors specifically designed for LGPR. In our experiments, we compared our proposed method with the Engineered method [7] and competitive handicraft descriptors (gprHOG [34] and HOVPS [1]) commonly used in underground perception tasks. The feature improved by the sequence-based approach

[1]Training Dataset from GROUNDED: *run:0008,0046,0047,0080*

[2]Test Dataset from GROUNDED: *run:0054,0056,0057,0018,0090*

TABLE II

COMPARISON AGAINST DIFFERENT APPROACHES IN MULTIPLE CHALLENGES: RECALL@1/5/15. BASED ON THE BENCHMARK SPECIFICATION [9], PERFORMANCE IN DIFFERENT WEATHERS USING ONLY MAPS COLLECTED ON SUNNY DAYS (MAP/QUERY)

| **Method/Challenge** | | **GROUNDED** | | | | **Self Dataset** | |
|---|---|---|---|---|---|---|---|
| | | Sunny/Sunny | Sunny/Snowy | Sunny/Rainy | Multi-Lane | Sunny/Sunny | Sunny/Rainy |
| *Single Scan:* | Engineered | 0.78/0.90/0.96 | 0.15/0.31/0.47 | 0.17/0.43/0.69 | 0.08/0.28/0.58 | 0.87/0.94/0.98 | 0.30/0.59/0.73 |
| | gprHOG | 0.76/0.88/0.96 | 0.09/0.23/0.43 | 0.16/0.41/0.72 | 0.08/0.29/0.57 | 0.79/0.89/0.95 | 0.19/0.49/0.76 |
| | HOVPS | 0.75/0.88/0.94 | 0.06/0.18/0.35 | 0.17/0.40/0.56 | 0.07/0.30/**0.59** | 0.86/0.95/0.98 | 0.28/0.61/0.84 |
| | MWS (Ours) | 0.77/0.88/0.94 | 0.14/0.33/0.48 | 0.06/0.25/0.51 | 0.06/0.25/0.50 | 0.89/0.96/0.98 | 0.39/0.61/0.87 |
| *Sequence Scans:* | seqSLAM | 0.02/0.90/0.99 | 0.01/0.37/0.63 | 0.02/0.02/0.65 | 0.02/0.03/0.47 | 0.02/0.59/0.90 | 0.01/0.66/0.93 |
| | Engineered_seq | 0.94/0.96/0.99 | 0.53/0.67/0.77 | 0.41/0.58/0.76 | 0.07/0.17/0.35 | **0.96/0.98/0.99** | 0.59/0.78/0.95 |
| | gprHOG_seq | 0.94/0.97/0.98 | **0.61**/0.72/0.81 | 0.35/0.50/0.66 | 0.08/0.19/0.33 | 0.94/0.97/0.98 | 0.45/0.65/0.79 |
| | HOVPS_seq | 0.95/0.97/0.99 | 0.58/0.72/0.81 | 0.39/0.53/0.68 | 0.16/0.30/0.45 | 0.94/0.97/**0.99** | **0.69**/0.81/**0.97** |
| | MWS_seq (Ours) | **0.97/0.99/1.00** | **0.61/0.76/0.86** | **0.57/0.76/0.89** | **0.17/0.38/0.59** | **0.96/0.98/0.99** | 0.67/**0.86**/0.95 |

TABLE III

ABLATION EXPERIMENTS OF SEQUENCE LENGTH SELECTION AND RERANKING: RECALL@1/5/15

| **Method/Dataset** | | **GROUNDED** | | | **Self Dataset** | | |
|---|---|---|---|---|---|---|---|
| | | L=5 | L=15 | L=25 | L=5 | L=10 | L=15 |
| *No Reranking:* | Engineered_seq | 0.44/0.63/0.73 | 0.69/0.82/0.87 | 0.81/0.88/0.93 | 0.55/0.70/0.84 | 0.67/0.78/**0.90** | 0.80/**0.89**/0.93 |
| | gprHOG_seq | 0.49/0.65/0.74 | 0.72/0.81/0.84 | 0.77/0.83/0.86 | 0.44/0.63/0.79 | 0.46/0.66/0.81 | 0.54/0.71/0.89 |
| | HOVPS_seq | 0.46/0.66/0.76 | 0.70/0.82/0.86 | 0.83/0.91/0.94 | 0.46/0.68/0.84 | 0.58/0.72/0.85 | 0.65/0.81/0.90 |
| | MWS_seq (Ours) | **0.51/0.71/0.80** | **0.79/0.89/0.91** | **0.90/0.94/0.96** | **0.62/0.74/0.87** | **0.72/0.81**/0.89 | **0.83**/0.87/**0.94** |
| *Reranking:* | Engineered_seq | 0.46/0.64/0.74 | 0.75/0.83/0.89 | 0.84/0.90/0.94 | 0.65/0.73/0.87 | **0.83**/0.88/0.96 | **0.91/0.94**/0.96 |
| | gprHOG_seq | **0.53/0.74/0.84** | 0.74/0.85/0.92 | 0.82/0.88/0.92 | 0.59/0.70/0.82 | 0.79/**0.92/0.98** | 0.80/0.93/**0.99** |
| | HOVPS_seq | 0.52/0.68/0.78 | 0.76/0.83/0.89 | 0.87/**0.94**/0.96 | 0.60/0.74/**0.89** | 0.75/0.85/0.93 | 0.83/0.89/0.94 |
| | MWS_seq (Ours) | 0.51/0.69/0.81 | **0.83/0.90/0.94** | **0.91/0.94/0.97** | **0.68/0.76**/0.88 | **0.83**/0.87/0.96 | 0.84/0.88/0.96 |

appends 'seq' to the suffix. The benchmark for sequence matching is SeqSLAM [19]. Table II presents quantitative results for different algorithms with a sequence length of 10 in various challenging environments, along with the Recall@K curve.

For all datasets, it can be observed that seqSLAM as a sequence-matching baseline achieves poor results considering Recall@K (1,5). We discovered that this result is due to the fact that the GPR scans in the dataset are captured closely with each other and, as a result, the contrast enhancement step on the difference matrix in seqSLAM would degrade its performance, as we verified experimentally.

The proposed MWS_seq consistently achieves the highest recall rates in most challenging scenarios. The application of sequence information effectively enhances the performance of all features, and even raw GPR data performs well under various weather conditions.

It is worth noting that, compared to the challenge posed by varying weather conditions, the improvement in performance for multi-lane scenarios within our framework is relatively modest. In multi-lane challenges, the low overlap of the underground scene observed by GPR leads to poor recognition. We have identified this as the primary source of LGPR positioning errors, a conclusion not previously reached in prior studies.

### *D. Effective of Sequence Length and Reranking*

Table III illustrates the impact of sequence length and reranking mechanisms on the average recall performance of all features in multi-weather challenges. As expected, increasing the sequence length will improve the matching performance of all features. It is worth noting that when the sequence length exceeds 15 (3-4m in reality), without feature processing, it can also effectively overcome multi-weather challenges.

The proposed reranking mechanism has enhanced the performance of all methods. The effectiveness of reranking is closely linked to the accuracy of past trajectories, where more accurate positioning results in the past lead to larger continuity weights assigned to the correct area.

When comparing our Self dataset to the public GROUNDED dataset, all methods outperform the latter. This can be attributed to the ultra-wide bandwidth of our radar equipment, which facilitates GPR scanning to capture more discriminative underground texture details.

These findings are expected to offer valuable insights for future research endeavors, particularly with regard to the exploration of adaptive sequence length selection, akin to the concept of *keyframe* selection in SLAM.

## V. CONCLUSION

In this paper, we open-source a framework based on sequence matching to fill the gap in current LGPR research. Additionally, we introduce a trainable network architecture designed to extract robust underground features in various weather conditions. Our proposed framework comprises four components, each offering potential for improvement, evaluation, and guidance for future algorithm design.

Leveraging sequence information allows LGPR to observe more underground scenes with few redundancies. It demonstrates that incorporating sequence information is more meaningful for LGPR compared to visual localization tasks. Furthermore, LGPR performance can be significantly

improved with just a dozen GPR scans, rather than requiring hundreds of scans as in the past.

This study encourages researchers to reconsider existing LGPR pipelines and explore the benefits of deep learning. In the future, we plan to expand our work to include multi-degree freedom localization estimation while considering the internal connections between multi-scans.

## ACKNOWLEDGEMENTS

We thank the anonymous reviewers and AC for their valuable comments. Research presented here has been supported by the National Natural Science Foundation of China under Grant Number 62201603, the China Postdoctoral Science Foundation under Grant Number 2023TQ0088, the Postdoctoral Fellowship Program of CPSF under Grant Number GZC20233539, and Research Program of National University of Defense Technology under Grant Number ZK22-04.

## REFERENCES

[1] P. Zhang, L. Shen, T. Wen, X. Huang, and Q. Xin, "Vector Phase Symmetry for Stable Hyperbola Detection in Ground-Penetrating Radar Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.

[2] M. Moalla, H. Frigui, A. Karem, and A. Bouzid, "Application of Convolutional and Recurrent Neural Networks for Buried Threat Detection Using Ground Penetrating Radar Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–13, 2020.

[3] L. J. Latecki and R. Lakamper, "Shape similarity measure based on correspondence of visual parts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1185–1190, 2000, publisher: IEEE.

[4] T. Ort, I. Gilitschenski, and D. Rus, "GROUNDED: The Localizing Ground Penetrating Radar Evaluation Dataset." in *Robotics: Science and Systems*, vol. 2, 2021.

[5] D. Merkle, C. Frey, and A. Reiterer, "ground penetrating radar and laser scanning for infrastructure mapping," *Journal of Applied Geodesy*, vol. 15, no. 1, pp. 31–45, 2021.

[6] Cornick, Koechling, Stanley, Zhang, and BJ, "Localizing Ground Penetrating RADAR: A Step Toward Robust Autonomous Ground Vehicle Localization," *J FIELD ROBOT*, vol. 2016,33(1), no. -, pp. 82–102, 2016.

[7] T. Ort, I. Gilitschenski, and D. Rus, "Autonomous navigation in inclement weather based on a localizing ground penetrating radar," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3267–3274, 2020.

[8] E. Skartados, A. Kargakos, E. Tsiogas, I. Kostavelis, D. Giakoumis, and D. Tzovaras, "Gpr antenna localization based on a-scans," in *2019 27th European Signal Processing Conference (EUSIPCO)*. IEEE, 2019, pp. 1–5.

[9] T. Ort, I. Gilitschenski, and D. Rus, "GROUNDED: A localizing ground penetrating radar evaluation dataset for learning to localize in inclement weather," *The International Journal of Robotics Research*, p. 02783649231183460, 2023.

[10] B. Bi, L. Shen, P. Zhang, X. Huang, Q. Xin, and T. Jin, "TSVR-Net: An End-to-End Ground-Penetrating Radar Images Registration and Location Network," *Remote Sensing*, vol. 15, no. 13, p. 3428, 2023.

[11] S. Wickramanayake, K. Thiyagarajan, and S. Kodagoda, "Deep learned ground penetrating radar subsurface features for robot localization," in *2022 IEEE Sensors*. IEEE, 2022, pp. 1–4.

[12] E. Hoxha, J. Feng, D. Sanakov, and J. Xiao, "Robotic Inspection and Subsurface Defect Mapping Using Impact-echo and Ground Penetrating Radar," *IEEE Robotics and Automation Letters*, 2023.

[13] A. Baikovitz, P. Sodhi, M. Dille, and M. Kaess, "Ground encoding: Learned factor graph-based models for localizing ground penetrating radar," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 5476–5483.

[14] S. Sapai, J. Y. Loo, Z. Y. Ding, C. P. Tan, R. C.-W. Phan, V. M. Baskaran, and S. G. Nurzaman, "Cross-domain transfer learning and state inference for soft robots via a semi-supervised sequential variational bayes framework," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 552–559.

[15] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "NetVLAD: CNN architecture for weakly supervised place recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5297–5307.

[16] S. Hausler, S. Garg, M. Xu, M. Milford, and T. Fischer, "Patch-netvlad: Multi-scale fusion of locally-global descriptors for place recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 14 141–14 152.

[17] H. Zhang, X. Chen, H. Jing, Y. Zheng, Y. Wu, and C. Jin, "Etr: An efficient transformer for re-ranking in visual place recognition," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 5665–5674.

[18] Z. Li, C. D. W. Lee, B. X. L. Tung, Z. Huang, D. Rus, and M. H. Ang, "Hot-NetVLAD: Learning Discriminatory Key Points for Visual Place Recognition," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 974–980, 2023, publisher: IEEE.

[19] M. J. Milford and G. F. Wyeth, "SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights," in *2012 IEEE international conference on robotics and automation*. IEEE, 2012, pp. 1643–1649.

[20] S. Garg, M. Vankadari, and M. Milford, "SeqMatchNet: Contrastive learning with sequence matching for place recognition & relocalization," in *Conference on Robot Learning*. PMLR, 2022, pp. 429–443.

[21] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[22] R. Mereu, G. Trivigno, G. Berton, C. Masone, and B. Caputo, "Learning sequential descriptors for sequence-based visual place recognition," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10 383–10 390, 2022, publisher: IEEE.

[23] E. Temlioglu and I. Erer, "A Novel Convolutional Autoencoder-Based Clutter Removal Method for Buried Threat Detection in Ground-Penetrating Radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[24] Z. Zhu, K. Lin, A. K. Jain, and J. Zhou, "Transfer learning in deep reinforcement learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, publisher: IEEE.

[25] H.-H. Sun, W. Cheng, and Z. Fan, "Learning to Remove Clutter in Real-World GPR Images Using Hybrid Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.

[26] Q. Dai, Y. He, Y. Lei, J. Lei, X. Wang, and B. Zhang, "GPR Data Reconstruction Using Residual Feature Distillation Block U-Net," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.

[27] S. Li, Q. Liu, and R. Li, "Buried object detection from GPR images using improved U-net," vol. 12552. SPIE, 2023, pp. 613–619.

[28] S. Dong, J. Jiao, S. Zhou, P. Lu, and Z. Zeng, "3-d gravity data inversion based on enhanced dual u-net framework," *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

[29] S. Zhu, L. Yang, C. Chen, M. Shah, X. Shen, and H. Wang, "R2former: Unified retrieval and reranking transformer for place recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 19 370–19 380.

[30] F. Tan, J. Yuan, and V. Ordonez, "Instance-level image retrieval using reranking transformers," in *proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 12 105–12 115.

[31] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981, publisher: ACM New York, NY, USA.

[32] P. Newbold, "Arima model building and the time series analysis approach to forecasting," *Journal of forecasting*, vol. 2, no. 1, pp. 23–35, 1983.

[33] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.

[34] D. Reichman, L. M. Collins, and J. M. Malof, "gprhog and the popularity of histogram of oriented gradients (hog) for buried threat detection in ground-penetrating radar," *arXiv preprint arXiv:1806.01349*, 2018.