SEBHA UNIVERSITY CONFERENCE PROCEEDINGS, VOL.5 NO. 4 2026
DOI: 10.51984/MTGHD007

وقائع مؤتمرات جامعة سبها
Sebha University Conference Proceedings
Confrence Proceeeding homepage: http://www.sebhau.edu.ly/journal/CAS

Deep Learning for Reliable RIS-Enabled SISO Localization Under Practical Model Mismatch

Amhimmid Alhashani*, Muamer Hawej

Department of Electrical and Computer Engineering, Faculty of Engineering, Al-Marqab University, Al-Khums , Libya

Keywords:
Reconfigurable intelligent surface (RIS)
localization
SISO
deep learning
robustness
model mismatch
reliability
p90

ABSTRACT

Reconfigurable Intelligent Surfaces (RISs) have shown strong potential for improving user localization in single-input single-output (SISO) systems, where both the base station (BS) and the user equipment (UE) are equipped with a single antenna. However, in practical wireless environments, localization accuracy can be affected by signal-to-noise ratio (SNR) variations, multipath propagation, and hardware-related impairments. These factors may create a mismatch between the assumed system model and the received signals, thereby reducing the performance of conventional model-based localization estimators. To address this issue, this work proposes a mixed-domain deep learning approach for three-dimensional downlink localization in RIS-enabled SISO systems. The proposed method constructs a learning observation tensor from OFDM pilot measurements by transforming the received signals into the delay domain using the inverse fast Fourier transform (IFFT), followed by appropriate normalization. The network is trained under different SNR levels and impairment conditions to improve its robustness against practical model mismatches. Simulation results show that the proposed method improves localization accuracy, as measured by root mean square error (RMSE), and enhances reliability, as measured by the 90th percentile error (p90). Compared with conventional model-based estimators, the proposed approach provides more stable performance across different SNR values and multipath scenarios, particularly in reducing the number of high-error cases.

التعلم العميق لتحديد موقع SISO الموثوق به باستخدام RIS في ظل عدم تطابق النموذج العملي

امحمد الحشاني* ومعمر حويج

قسم الهندسة الكهربائية و الحاسوب ، كلية الهندسة ، جامعة المرقب ، الخمس ، ليبيا

الكلمات المفتاحية:
سطح ذكي قابل لإعادة التشكيل (RIS)
تحديد الموقع
مدخل واحد، مخرج واحد (SISO)
التعلم العميق
المتانة
عدم تطابق النموذج
الموثوقية
النسبة المئوية 90% (P90)

الملخص

أظهرت الأسطح الذكية القابلة لإعادة التشكيل (RISs) إمكانات كبيرة لتحسين تحديد موقع المستخدم في أنظمة الإدخال والإخراج الأحادي (SISO)، حيث تُجهز كل من محطة القاعدة (BS) وجهاز المستخدم (UE) بهوائي واحد. مع ذلك، في بيئات الاتصالات اللاسلكية العملية، قد تتأثر دقة تحديد الموقع بتغيرات نسبة الإشارة إلى الضوضاء (SNR)، وانتشار الإشارات عبر مسارات متعددة، والتشوهات المتعلقة بالأجهزة. قد تُحدث هذه العوامل عدم تطابق بين نموذج النظام المفترض والإشارات المُستقبلة، مما يُقلل من أداء مُقدِّرات تحديد الموقع التقليدية القائمة على النموذج. لمعالجة هذه المشكلة، يقترح هذا العمل نهجًا للتعلم العميق متعدد المجالات لتحديد الموقع ثلاثي الأبعاد في وصلة الإرسال الهابطة في أنظمة SISO المُفعَّلة بتقنية RIS. تُنشئ الطريقة المقترحة موترًا للملاحظة من قياسات إشارات OFDM التجريبية عن طريق تحويل الإشارات المُستقبلة إلى مجال التأخير باستخدام تحويل فوريه العكسي السريع (IFFT)، متبوعًا بتطبيع مناسب. يتم تدريب الشبكة في ظل مستويات مختلفة من نسبة الإشارة إلى الضوضاء وظروف تشوهات مختلفة لتحسين متانتها في مواجهة حالات عدم تطابق النموذج العملية. تُظهر نتائج المحاكاة أن الطريقة المقترحة تُحسّن دقة تحديد الموقع، كما يُقاس بمتوسط الجذر التربيعي للخطأ (RMSE)، وتعزز الموثوقية، كما يُقاس بخطأ النسبة المئوية التسعين (p90). وبالمقارنة مع المُقدِّرات التقليدية

*Corresponding author:
E-mail addresses: asdefg2175@gmail.com, (M. Hawej) m.malhwaij@elmergib.edu.ly
Article History : Received 30 May 2026 - Received in revised form 30 September 2026 - Accepted 11 November 2026

القائمة على النماذج، تُوفر الطريقة المقترحة أداءً أكثر استقرارًا عبر قيم نسبة الإشارة إلى الضوضاء المختلفة وسيناريوهات تعدد المسارات، لا سيما في تقليل عدد حالات الخطأ العالي.

I. INTRODUCTION

Reconfigurable Intelligent Surfaces (RISs) have emerged as a pivotal technology for sixth-generation (6G) wireless networks by dynamically shaping the radio propagation environment via programmable, energy-efficient metasurfaces capable of imposing controllable phase and amplitude modifications on incident electromagnetic waves [1]. Beyond enhancing coverage and mitigating signal blockages, RISs facilitate novel opportunities for precise radio localization by creating additional, structured non-line-of-sight (NLOS) propagation paths and directing signal energy toward designated spatial regions [2]. Simultaneously, machine learning (ML) techniques have advanced significantly as robust data-driven approaches for RF-based positioning, enabling inference from diverse signal features such as received signal strength, time and angle of arrival, and channel state information (CSI), while effectively compensating for hardware imperfections and dynamic environmental conditions [3]. Recent progress in angle-of-arrival (AOA) estimation has seen the adoption of deep learning (DL) techniques to effectively address challenges such as various system imperfections, low signal-to-noise ratio (SNR) scenarios, and multipath propagation, while also offering solutions with reduced computational complexity compared to traditional model-based methods [4].

Localization and synchronization can be achieved by estimating angles of arrival (AOAs), angles of departure (AODs), and delays, even when the multipath sources are unknown beforehand [5].In [6], the problem of 2D localization with (SISO), synchronized signaling, and multiple RISs equipped with uniform linear arrays (ULAs) is analyzed by deriving the Cramér-Rao Bound (CRB) limits. Position error bounds (PEBs) have been derived for 3D uplink localization considering both synchronous and asynchronous cases. It was noted that these bounds can be attained by applying a grid search to maximize the log-likelihood function; however, no low-complexity estimation algorithm was presented [7].

Recent work by Hassan et al. demonstrated that machine learning can support user localization in RIS-assisted wireless systems by leveraging received-power measurements from limited probing beams, thereby reducing beam-search overhead and improving localization efficiency in future 6G networks [8]. Reconfigurable Intelligent Surfaces (RISs) have emerged as a fundamental technology driving cost-effective and energy-efficient strategies for optimizing wireless communication systems. By altering the properties of electromagnetic wave propagation within the environment, RISs facilitate substantial enhancements in wireless communication performance [9].

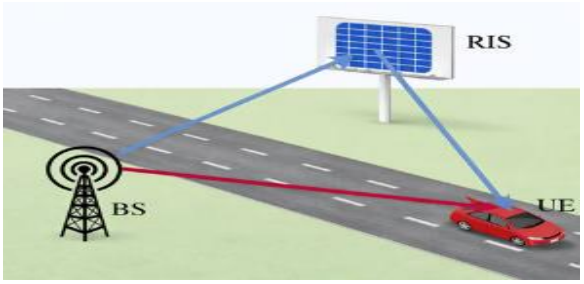
Furthermore, RIS-assisted localization techniques exploit the capability of RISs to actively control multipath propagation, thereby improving the accuracy and reliability of localization systems. This approach effectively extends the concept of multipath-aided localization to incorporate RIS technology [10].

Recent studies have explored RIS-enabled localization within simplified frameworks. Specifically, the fundamental SISO RIS-assisted model was presented in [11]. where joint localization and synchronization problems were thoroughly examined.

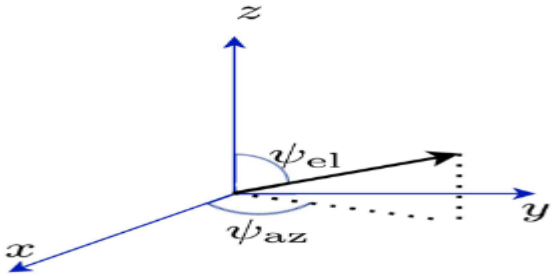
However, these model-based methods do not work well when there are real-world mismatches.

In this paper, we investigate 3D downlink localization and synchronization of a user equipment (UE) equipped with a single antenna, utilizing the received line-of-sight (LOS) signal from a single-antenna base station (BS) along with the reflected signal from a reconfigurable intelligent surface (RIS) featuring a uniform planar array (UPA).

We also propose a mixed-domain deep learning estimator for RIS-aided localization and synchronization, robust to practical mismatches. The proposed DL method improves both RMSE and reliability metrics compared to the classical mismatch baseline, as verified through simulations.



(a) Fig. 1: (a) System setup



(b) Elevation and azimuth angles of a vector.

In this study, we demonstrate that the upcoming 6G technology can leverage new enablers, including reconfigurable intelligent surfaces (RIS), to accurately estimate the user equipment (UE) position, clock bias, and velocity, even in single-input single-output (SISO) wireless link scenarios.

II. System Model

This section introduces the considered RIS-assisted SISO localization system and the learning-ready observation used by the proposed method. The system geometry and physical measurement model follow the baseline formulation in [11, Sec. II]. During repeated pilot transmissions, the RIS applies different phase-shift profiles, allowing the UE to observe both the direct BS–UE path and the reflected BS–RIS–UE path under different configurations. The collected measurements are then transformed into a robust observation tensor and used as input to the proposed deep learning framework.

A. System Geometry and RIS Protocol

We examine a three-dimensional setting involving a single-antenna base station (BS) and a single-antenna user equipment (UE), supported by a reconfigurable intelligent surface (RIS). The RIS consists of a uniform planar array (UPA) with $\boldsymbol{M_r} \times \boldsymbol{M_C}$ reflecting elements. The BS position $\mathbf{p}_{\boldsymbol{b}} \in \mathbb{R}^{\boldsymbol{3}}$and RIS position $\mathbf{p}_{\boldsymbol{r}} \in \mathbb{R}^{\boldsymbol{3}}$ are assumed known, whereas the UE position $\boldsymbol{p} \in \boldsymbol{R^3}$ is unknown and to be estimated.

$$\boldsymbol{R} \triangleq [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3]^{\mathsf{T}} \tag{1}$$

where $\boldsymbol{R}$ is the RIS rotation matrix, and $\boldsymbol{v_1}, \boldsymbol{v_2}$ and $\boldsymbol{v_3}$are orthonormal vectors defining the local coordinate system of the RIS.

To generate measurement diversity within the basic SISO configuration. The RIS applies T distinct phase profiles across repeated pilot transmissions. Let $\boldsymbol{Y}_t$ denote the vector of complex unit-modulus RIS coefficients at transmission t, and let diag $(\boldsymbol{Y}_t)$ be the corresponding diagonal phase-shift matrix, for $\boldsymbol{t = 0, \ldots, T - 1}$. Optionally, in the asynchronous scenario, we account

for an unknown UE clock bias Δ_t which appears as an extra timing offset in the received pilot signals.

B. Signal/Measurement Model

The BS transmits known pilots over. $\boldsymbol{N_s}$ OFDM subcarriers. For each RIS phase profile $\boldsymbol{t}$, the user equipment gathers the matching frequency-domain pilot samples. Stacking these observations across the $\boldsymbol{T}$ profiles yield the measurement matrix $\mathbf{Y} \in \mathbb{C}^{\boldsymbol{N_s \times T}}$. Under the baseline model [11], the received signal is the superposition of the direct LOS component and the RIS-reflected component, plus additive noise:

$$\boldsymbol{Y} = g_{\mathrm{b}}\sqrt{E_{\mathrm{s}}}\boldsymbol{d}(\tau_{\mathrm{b}})\mathbf{1}_T^{\mathrm{T}} + g_{\mathrm{r}}\sqrt{E_{\mathrm{s}}}\boldsymbol{d}(\tau_{\mathrm{r}})\boldsymbol{u}(\boldsymbol{\phi})^{\top} + \boldsymbol{N}, \qquad (2)$$

Here, $\boldsymbol{g_{\mathrm{b}}}$ and $\boldsymbol{g_{\mathrm{r}}}$ denote the complex path gains, $\boldsymbol{\tau_{\mathrm{b}}}$and $\boldsymbol{\tau_{\mathrm{r}}}$ denote the corresponding propagation delays, and $\boldsymbol{E}$ is the transmit signal energy. The function $\boldsymbol{d}(.)$ represents the delay response across subcarriers, and $\boldsymbol{u}(\boldsymbol{\phi})$ denotes the RIS-dependent response. In the asynchronous case, the clock bias $\boldsymbol{\Delta_t}$ is absorbed into the effective delays. The vector 1 is an all-ones vector of length **T**, and $\boldsymbol{N}$ is an additive noise matrix. The delay response across subcarriers is defined as [11].

$$\boldsymbol{d}(\tau) = \left[1, e^{-j2\pi\tau\Delta f}, \ldots, e^{-j2\pi\tau(N-1)\Delta f}\right]^{\mathrm{T}} \qquad (3)$$

where $\boldsymbol{d}(\boldsymbol{\tau})$ represents the delay response vector across subcarriers, $\boldsymbol{\Delta f}$ is the subcarrier spacing, and $\boldsymbol{N_s}$is the number of subcarriers. The RIS-dependent response is expressed as [11].

$$\begin{aligned}[\boldsymbol{u}(\phi)]_t &= (\boldsymbol{a}(\boldsymbol{\theta}))^{\mathrm{T}}\mathrm{diag}(\boldsymbol{\gamma}_t)\boldsymbol{a}(\phi) \\ &= (\boldsymbol{\gamma}_t^{\mathrm{T}} \odot (\boldsymbol{a}(\boldsymbol{\theta}))^{\mathrm{T}})\boldsymbol{a}(\phi)\end{aligned} \qquad (4)$$

where $[\boldsymbol{u}(\boldsymbol{\phi})]_{\boldsymbol{t}}$ denotes the ***t-th*** element of the RIS response vector $\boldsymbol{\gamma}_t \in \mathbb{C}^M$ is the RIS phase shift vector at transmission $\boldsymbol{t}$ **,** $\mathbf{diag}(\boldsymbol{\gamma_t})$is the diagonal matrix of RIS phase shifts, and $\mathbf{a}(\boldsymbol{\theta})$ and $\mathbf{a}(\boldsymbol{\phi})$ are the array response vectors corresponding to the angle-of-arrival (**AOA**) $\boldsymbol{\theta}$ and angle-of-departure (**AOD**)$\boldsymbol{\phi}$, respectively. This term captures the effect of RIS phase configurations and array geometry on the received signal.

C. Learning Observation Tensor

While the physical model is naturally expressed in the frequency domain, we construct a compact input representation for learning. Specifically, we (optionally) select/down sample the subcarriers in a consistent manner to obtain $\mathbf{Y}_{ds}$, then transform the observations to the delay domain via an IFFT and apply a log-magnitude mapping to stabilize dynamic range:

$$\mathrm{X} = \log\left(\left|\,\mathrm{IFFT}\{\ \mathbf{Y}_{ds}\}\right| + \varepsilon\right),\ \ (\varepsilon > 0) \qquad (5)$$

where X is the resulting observation tensor used as input to the deep learning model, $\mathbf{Y_{ds}} \in \mathbb{C}^{\boldsymbol{N_s \times T}}$ is the (optionally) downsampled received signal matrix, IFFT(·) denotes the inverse fast Fourier transform applied along the subcarrier dimension, and $\boldsymbol{\varepsilon > 0}$ is a small constant introduced to ensure numerical stability.

The resulting delay-domain representations are stacked across the **T** RIS phase profiles to form the observation tensor **X**. In our implementation, each sample is represented as a tensor of size

1 × N_delay × T, and X is normalized using dataset statistics before training and inference

This change makes the signal's time-delay features clearer and helps the system remain robust even when noise or significant changes in signal strength occur.

D. Measurement Metrics

In real life, differences between the model and actual conditions, along with hardware issues, can make traditional estimation methods perform poorly. This is why using data-based methods is a good idea. Given the observation tensor X, our objective is to estimate the UE state. In the main setting, this corresponds to estimating the **3D** UE position p (and optionally the clock bias Δ_t). We learn a regression mapping $\hat{\mathbf{p}} = f_\theta(\mathbf{X})$, where θ denotes the network parameters. The position error is defined as

$$\boldsymbol{e} = \|\hat{\mathbf{p}} - \mathbf{p}\|_2 \qquad (\mathbf{6})$$

where e denotes the localization error, $\hat{\mathrm{p}}$ is the estimated UE position, $\|\cdot\|_2$ denotes the Euclidean norm, and $f_\theta(\cdot)$ represents the deep learning model parameterized by θ.

We report the root-mean-square error (RMSE) and the 90th percentile error (p90) as reliability metrics. The p90 denotes the value below which 90% of localization errors fall, effectively capturing the estimator's tail performance and robustness to large errors.

Remark: Equations (1)–(4) follow the baseline geometry and measurement model in [8, Sec. II]; the proposed contribution in this paper is the learning observation tensor construction and the mixed-domain learning framework introduced in Sections IV–VI.

III. Baseline Classical Estimator (Reference)

As a baseline, we consider the classical likelihood-based localization and synchronization method proposed in [11]. Although an exhaustive grid search over the unknown parameters can theoretically approach the performance bounds, the method in [11] reduces the computational burden by decomposing the original multi-dimensional optimization problem into a set of lower-dimensional searches, including two one-dimensional searches and one two-dimensional search. This provides near-optimal localization performance with considerably lower complexity. However, in practical scenarios, the baseline relies on an assumed signal model. Non-ideal effects such as imperfect RIS phase control, frequency offsets, multipath mismatch, quantization errors, and phase noise can degrade the performance of the classical estimator, because the assumed likelihood function may no longer match the true measurement distribution. This limitation motivates the development of a data-driven, robust estimator that better handles practical model mismatches.

- Baseline refers to the classical estimator from [11], based on grid/maximum likelihood principles and low-complexity decomposition.

- It serves as a relevant reference because it is model-based and tied to the CRB/PEB analysis in [11].

- Model inaccuracies and impairments can significantly reduce accuracy and reliability, which motivates the need for deep learning-based robustness.

IV. Proposed Mixed-Domain Deep Learning Method

We will introduce a multi-domain deep learning estimator that directly predicts the user's state from the future tensor. This estimator treats the model (X) as input and produces (hat{Y} = f_theta(X)), where (theta)represents the parameters to be trained.

We use a convolutional neural network architecture specifically designed for coordinated measurement tensors.Many comprehensive, matching deep learning approaches have proven effective in wireless positioning in challenging deployment environments [12]. Surveys have confirmed its resilience [13].

Mixed-domain training. To enhance generalization ability under practical model mismatches the network is trained using a variety of conditions that include different levels of signal-to-noise ratio (SNR) and noise conditions. For example, the effects of multipath processing and technical flaws in devices. yields more reliable mappings. This model is better able to handle scenario variations. These factors often influence model-based estimators. During training. The regression loss function is applied. The mean squared error scale is applied. The mean squared error (MSE) scale is used to analyze the regression. Input data is also normalized using standardization or rationing for each channel. and the dataset is divided into training and testing sets,

validation. In addition, early stopping and learning rate scheduling are used to reduce overfitting and support stable training.

• model: $\hat{Y} = f_\theta$ (X) (regression of 3D position and optional clock bias).
• mixed-domain training over SNR + multipath/impairments for robustness.
• Loss: MSE (or weighted MSE); normalization; train/val/test split.
• Optional: inference-time comparison vs classical grid/peak+grid.

V. Dataset Generation and Training Protocol

We are working on creating a supervised dataset using Monte Carlo simulation for a RIS-supported SISO system as described in [11]. For each sample , The(UE) device location is chosen randomly (if there is clock deviation). The corresponding received measurements are then generated based on a specific timetable for this RIS phase profile. With the addition of distortions and blurring, The input tensor (X) and regression (Y) are stored in the HDF5 dataset to facilitate efficient training.

Scenario design. Two main stress tests are performed. (i) SNR ratio survey,SNR changes across a range of values, while other information remains constant. (ii) In a multi-path scan, During this process, the multipath strength is modified to assess robustness under increased model mismatch conditions.

Training protocol. The dataset is divided into training, testing, and validation sets.(e.g 80/ 10 / 10). Training is performed using the Adam optimizer with mini-batches and either a fixed or scheduled learning rate over a moderate number of epochs. All hyperparameters remain consistent across experiments to ensure a fair comparison between single-domain and mixed-domain models.

• Dataset: (X, Y) from Monte Carlo simulation aligned with [1].
• Sweeps: SNR sweep and multipath sweep (mpRel Max).
•Train/val/test split; Adam; mini-batch training; same hyperparameters for fairness.

Table I. Simulation and dataset parameters.

Parameter	Value
Carrier frequency	30 GHz
Subcarrier spacing	120 kHz
Number of subcarriers (full)	3000
Selected subcarriers (feature depth)	128
RIS UPA size	64 × 64 elements (4096)
RIS phase profiles (full)	256
Selected phase profiles (feature width)	64
UE range/height	r ∈ [1, 35] m, z = −10 m
Reference SNR for CDF & MP sweep	10 dB
SNR sweep	0:5:30 dB
Multipath sweep	mpRel Max ∈ {0, 0.1, 0.2, 0.3, 0.4}
Training settings	Adam, LR=1e−3, batch=64, epochs=6
Evaluation	Mean over 3 random seeds

VI. Simulation Results

This part displays the outcomes of the conducted simulations, examining the performance of the proposed models under various test scenarios.

A. Cumulative Distribution Function at a Reference Setting (e.g., SNR = 10 dB)

First we give the cumulative distribution function (CDF) of the position error at a typical operating point e.g. SNR = 10 dB. The CDF tells us about the average accuracy as well as the tail behavior of the distribution. We compare three methods: (i) the classical mismatch baseline, (ii) The term Deep Learning Single-Domain refers to a neural network trained using data generated under a single operating condition and serves as the learning-based baseline for comparison., and (iii) The proposed mixed-domain deep learning model refers to a neural network trained using data generated under different operating conditions, including various SNR levels, multipath strengths, and system impairments.

The mixed-domain model shows a left-shifted CDF, indicating improved reliability especially in the higher percentiles.

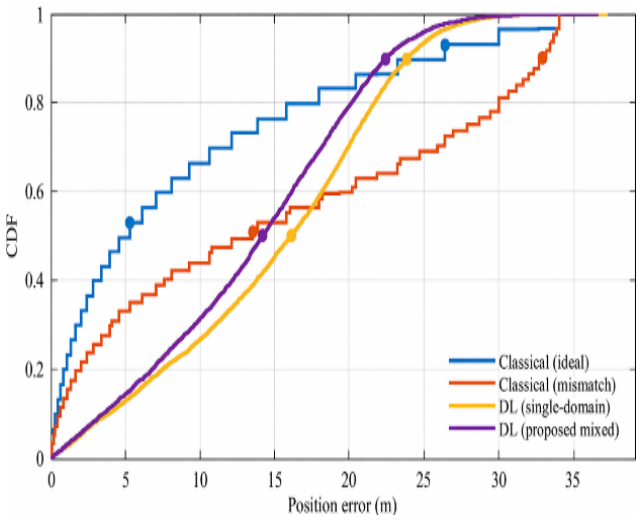


Fig. 2. Empirical CDF of the 3D position error at SNR=10 dB

At the reference scenario (SNR = 10 dB), the proposed mixed-domain model enhances both the accuracy and tail reliability compared to the classical estimator experiencing mismatch, as shown in (Table II).

Table II. Position error statistics

Method	RMSE (m)	Median (m)	p90 (m)
Classical (ideal)	13.109	5.290	26.389
Classical (mismatch)	19.659	13.566	32.913
DL (single-domain)	16.735	16.068	23.851
Proposed DL (mixed)	15.393	14.167	22.478

B. Performance vs SNR (RMSE and p90)

Next, we evaluate the performance over a range of SNR values. The root mean square error (RMSE) and the 90-th percentile error (p90) are plotted as a function of SNR for the different methods. In this range the mixed domain deep learning model consistently outperforms the classical mismatch baseline for both RMSE and p90 values. This indicates a better robustness when the noise level decreases and the performance of the estimator is constrained by the mismatch.

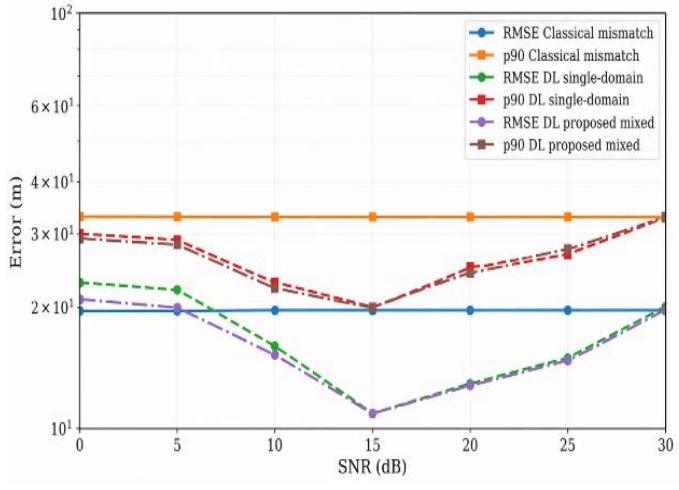


Fig. 3. RMSE (circles) and p90 (squares) versus SNR under classical mismatch and DL-based estimator

The proposed mixed-domain model is consistently better than the classical mismatch baseline across the entire SNR range. Typical RMSE improvements are of the order of 20-45% and typical p90 improvements are of the order of 25-40% at representative SNR points.

C. Robustness vs Multipath (RMSE and p90)

Finally, we test the estimators in the scenario of increasing multipath mismatch by varying the multipath strength parameter (mpRel Max) and reporting RMSE and p90. The proposed mixed-domain deep learning approach shows better robustness with much lower degradation in p90 as the multipath intensity increases.

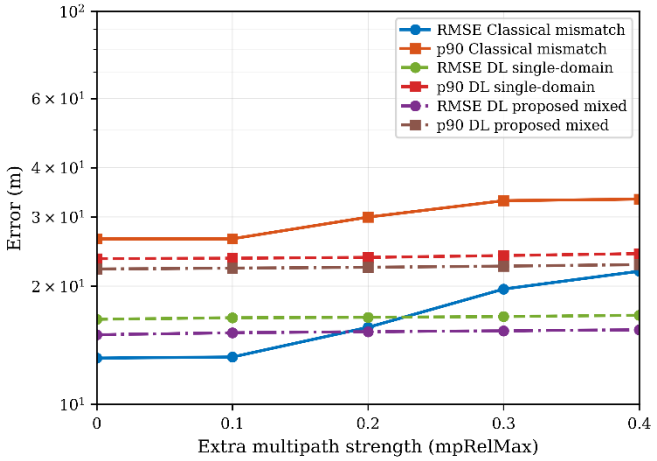


Fig. 4. RMSE (circles) and p90 (squares) versus multipath strength (mpRel Max) at SNR=10 dB

Figure 4 compares localization performance under increasing multipath mismatch. The Single-Domain model achieves lower RMSE and p90 than the classical mismatch estimator, confirming the effectiveness of deep learning for RIS-enabled localization. However, it remains less robust than the proposed Mixed-Domain model, particularly at higher multipath levels. This is because the Single-Domain model is trained under one operating condition, such as a fixed SNR and channel environment, whereas the Mixed-Domain model is trained across diverse SNR and multipath conditions, improving its generalization and robustness.

D. Summary Table (Gains)

Table III quantifies the improvement of the proposed mixed-domain model over the classical mismatch baseline.

• At SNR=10 dB, the proposed model reduces RMSE from 19.659 m to 15.393 m and p90 from 32.913 m to 22.478 m (Table II).

Table III. Relative gains (%) of proposed DL (mixed) versus classical mismatch.

Condition	RMSE gain (%)	p90 gain (%)
SNR = 10 dB	21.71	31.85
SNR = 15 dB	43.83	39.15
SNR = 20 dB	33.92	25.00
mpRel Max = 0.2	2.51	25.45
mpRel Max = 0.4	29.03	31.87

• Across different SNR levels, the improvements remain steady, with the most significant gains observed at moderate SNR values (see Table III).

• Under strong multipath (mpRel Max=0.4), the proposed model improves both RMSE (+29.03%) and p90 (+31.87%) versus mismatch (Table III).

E. Computational Efficiency

The proposed model contains 258.371 trainable parameters, requires approximately 1.033 MB for FP32 parameter storage, and achieves an average CPU inference time of 3.06 ms per sample using a batch size of one. The peak additional MATLAB process memory during inference was 67.75 ± 0.23 MB. Measurements were conducted using MATLAB R2021a on an Intel Core i7-10750H CPU.

VII. Conclusion

In this paper, robust three-dimensional user localization in an RIS-enabled SISO system was investigated under practical model-mismatch conditions. Building on a recently proposed RIS-assisted localization framework, a mixed-domain deep learning approach was developed to enhance localization performance and improve robustness against variations in SNR and multipath propagation.

The proposed method converts the received signal measurements into a structured learning observation tensor, which is then used as input to a supervised deep learning model for estimating the user position. The simulation results show that the proposed mixed domain learning approach performs better than the classical mismatch-based estimator and the single domain learning model. This is reflected in terms of RMSE and p90 localization error metrics. In particular, the reduction in the p90 error for the proposed method suggests greater reliability by reducing the probability of large localization errors in practical operating conditions.

Overall, the obtained results confirm that mixed-domain deep learning is an effective and robust solution to improve the accuracy and reliability of RIS-enabled localization systems in the presence of model mismatch. Future directions include validation of the proposed approach with real measurement data, extension of the framework to RIS-enabled MIMO localization systems, investigation of more complicated NLOS propagation environments, and design of lightweight neural network architectures for real-time implementation.

References

[1] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead,” IEEE Journal on Selected Areas in Communications, vol. 38, no. 11, pp. 2450–2525, 2020.
[2] A. Umer, I. Mu¨ursepp, M. M. Alam, and H. Wymeersch, “Role of ¨ reconfigurable intelligent surfaces in 6G radio localization: Recent developments, opportunities, challenges, and applications,” arXiv:2312.07288, 2024.
[3] D. Burghal, A. T. Ravi, V. Rao, A. A. Alghafis, and A. F. Molisch, “A comprehensive survey of machine learning based localization with wireless signals,” arXiv:2012.11171, 2020.
[4] Y. Liu, I. Al-Nahhal, O. A. Dobre, and F. Wang, “Deep-learning channel estimation for IRS-assisted integrated sensing and communication system,” IEEE Trans. Veh. Technol., vol. 72, no. 5, pp. 6181–6193, May 2023.
[5] H. Wymeersch,N.Garcia,H.Kim,G.Seco-Granados,S.Kim,F.Wen, and M .Fr¨ohle, "5Gmmwavedownlinkvehicularpositioning," inIEEE Global Commun .Conf. (GLOBECOM).,AbuDhabi,UAE,Dec. 2018, pp.206–212.
[6] H. WymeerschandB.Denis, "Beyond5Gwireless localizationwith reconfigurableintelligentsurfaces," inIEEEInt.Conf.Commun. (ICC), Dublin, Ireland, Jun.2020.
[7] A Elzanaty,A.Guerra,F.Guidi,andM.-S.Alouini,“Reconfigurablein telligentsurfacesforlocalization:Positionandorientationerrorbounds,” arXivpreprintarXiv:2009.02818,2020.
[8] M. T. Hassan, D. Zelenchuk, and M. A. B. Abbasi, "Machine Learning-Driven User Localization in RIS-Assisted Wireless Systems," arXiv:2510.23908, 2025.
[9] G. C. Alexandropoulos, I. Vinieratou, and H. Wymeersch, “Localization via multiple reconfigurable intelligent surfaces equipped with single receive RF chains,” IEEE Wireless Commun. Lett., vol. 11, no. 5, pp. 1072–1076, May 2022.
[10] A. Fascista, M. F. Keskin, A. Coluccia, H. Wymeersch, and G. Seco-Granados, “RIS-aided joint localization and synchronization with a single-antenna receiver: Beamforming design and

lowcomplexity estimation," IEEE J. Sel. Topics Signal Process., vol. 16, no. 5, pp. 1141–1156, Aug. 2022.
[11] K. Keykhosravi, M. F. Keskin, G. Seco-Granados, and H. Wymeersch, "SISO RIS-Enabled Joint 3D Downlink Localization and Synchronization," IEEE, 2021.
[12] R. Ayyalasomayajula et al., "Deep Learning Based Wireless Localization for Indoor Navigation," in Proc. ACM Mobi Com, 2020, doi:10.1145/3372224.3380894.
[13] K. A. Kordi et al., "Survey of Indoor Localization Based on Deep Learning," , doi:10.32604/cmc.2024.044890.2024.