Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sparse Representation and Multi-Stage Adaptive Filtering for Low-Velocity GMTI in SAR Imagery

Ashkan Masoomi
Islamic Azad University of Shiraz
Naser Parhizgar
n.parhizgar@iau.ac.ir
Islamic Azad University of Shiraz
Bijan Abbasi arand
Tarbiat Modares University

Research Article

Keywords:

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11679294/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Sparse Representation and Multi-Stage Adaptive Filtering for Low-Velocity GMTI in SAR Imagery

Ashkan Masoomi[1], Naser Parhizgar[1*], Bjan Abbasi Arand[2]

[1]Department of Electrical Engineering, SH.C., Islamic Azad University, Shiraz, Iran.
[2]Department of Electrical Engineering, Tarbiat Modares University, Tehran, Iran.

*Corresponding author(s). E-mail(s): n.parhizkar@iau.ac.ir;

Abstract

Detecting small and low-velocity targets in synthetic aperture radar (SAR) imagery remains a critical challenge for ground-moving target indication (GMTI), primarily because their weak signatures are often masked by dominant clutter. Conventional methods rely on partial signal features, leading to insufficient discrimination. This paper proposes a synergistic framework that integrates sparse representation with multi-stage adaptive filtering. First, a sparse model estimates and suppresses dominant clutter to produce a residual representation. Subsequently, a hierarchical filter combines Doppler, phase, and spatial-temporal features to enhance target discrimination. Evaluated on a simulated dataset replicating TerraSAR-X dual-receive antenna (DRA) along-track interferometry (ATI) parameters, the framework is tested across urban, suburban, and rural environments. Compared against displaced phase center antenna (DPCA), space-time adaptive processing (STAP), and sparse Bayesian learning (SBL) baselines, the proposed cascade achieves a detection probability of 0.76 at a false-alarm rate of 10^{-3}, and a minimum detectable velocity (MDV) of 1.2 m/s, which is 1.4 times the theoretical Doppler resolution limit. A simulated case study on urban traffic monitoring demonstrates its operational potential, achieving a mean velocity error of 0.10 m/s. With a computational complexity of $\boldsymbol{O(LKT + L^3)}$, the framework offers a computationally efficient approach for potential applications in security, environmental monitoring, and intelligent surveillance.

Keywords: Synthetic aperture radar (SAR); ground-moving target indication (GMTI); along-track interferometry (ATI); sparse representation; multi-stage adaptive filtering; low-velocity target detection.

1 Introduction

As a pivotal Earth observation tool, synthetic aperture radar (SAR) provides high-resolution imaging regardless of illumination or weather conditions [1], [2]. Beyond static scene reconstruction, SAR supports ground-moving target indication (GMTI), particularly for traffic monitoring [2]–[4]. However, detecting small and low-velocity targets remains challenging because their weak radar returns are frequently masked by dominant ground clutter [5]–[8].

The detectability of moving targets depends on radar geometry, target kinematics, and background scattering [9]–[11]. Depending on the sensor configuration, target motion may appear as Doppler shifts, interferometric phase changes, or spatial displacement [2], [8], [11]–[13]. For low-velocity targets, these signatures are often too subtle to distinguish from stationary clutter. Their

small spatial extent further increases the likelihood of masking, while reliance on a single feature or static threshold can lead to missed detections and false alarms [7], [9], [12], [14].

Existing GMTI methods exploit different signal characteristics. Displaced phase center antenna (DPCA) suppresses stationary returns through channel subtraction but requires accurate platform stabilization [2], [8], [15]. Space-time adaptive processing (STAP) provides effective clutter rejection but depends on sufficient homogeneous training data [5], [16]. Along-track interferometry (ATI) exploits inter-channel phase differences and is suitable for traffic monitoring, although calibration errors and temporal decorrelation can degrade its performance [2], [11], [12], [17]. Sparse Bayesian learning (SBL) jointly estimates clutter and target components but requires careful hyperparameter tuning [16], [18], [19]. Although sparse representation can model dominant clutter [7], [9], [20], [21], its residual may still contain unmodeled scatterers. Similarly, a single adaptive filter may be insufficient when target features vary across heterogeneous environments.

These limitations motivate a unified architecture combining sparse clutter modeling with complementary Doppler, ATI-phase, and spatial-temporal filtering stages. To address this need, we propose a multi-stage framework for low-velocity GMTI using simulated TerraSAR-X-like dual-receive-antenna (DRA) data in an ATI configuration [1], [2]. Sparse representation first estimates the dominant clutter and produces a residual signal for subsequent analysis [20], [21]. The residual is then processed through hierarchical filtering stages that exploit complementary target features to improve discrimination while controlling false alarms [9], [14]. The framework is evaluated across urban, suburban, and rural clutter environments.

The key contributions of this work are threefold. First, a sparse-representation-based mechanism is integrated into the detection architecture for dominant clutter suppression. Second, a hierarchical filtering strategy combines complementary Doppler, ATI-phase, and spatial-temporal information to improve target discrimination. Third, the proposed framework is quantitatively compared with ATI, DPCA, STAP, and SBL baselines in terms of detection probability, minimum detectable velocity (MDV), and computational cost.

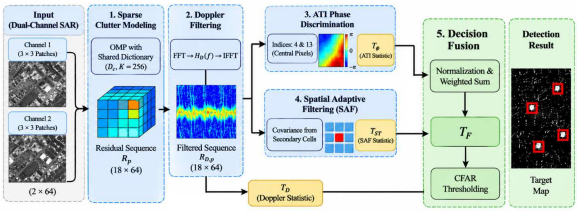

Fig. 1 Block diagram of the proposed sparse representation and multi-stage adaptive filtering framework for low-velocity GMTI in SAR imagery. The framework consists of four sequential stages: sparse clutter modeling, Doppler filtering, ATI phase discrimination, and spatial-temporal filtering, followed by decision fusion with CFAR thresholding.

The remainder of this paper is organized as follows. Section 2 describes the proposed methodology, Section 3 presents the experimental design and results, and Section 4 concludes the paper.

2 Proposed Method

The proposed framework processes TerraSAR-X-like dual-receive-antenna (DRA) data acquired in an along-track interferometric (ATI) configuration [1], [2]. It combines sparse clutter modeling, multi-stage adaptive filtering, and decision fusion to exploit spatial, Doppler, and inter-channel phase information for detecting moving targets against stationary clutter [2], [8], [12]. The overall architecture is illustrated in Fig. 1.

2.1 Data Model and Sparse Clutter Suppression

Let $z_1(p,m)$ and $z_2(p,m)$ denote the complex-valued observations of resolution cell p, acquired by the fore and aft receive channels, respectively, at pulse m. A local observation vector is constructed by concatenating the two corresponding 3×3 spatial patches using a column-major vectorization operation:

$$\mathbf{y}_{p,m} = \begin{bmatrix} vec\,(\mathcal{P}_{1,p,m}) \\ vec\,(\mathcal{P}_{2,p,m}) \end{bmatrix} \in \mathbb{C}^{18} \tag{1}$$

where $\mathcal{P}_{1,p,m}$ and $\mathcal{P}_{2,p,m}$ represent the local 3×3 patches centered at resolution cell p and pulse m from the fore and aft receive channels, respectively. The observation can be expressed as

$$\mathbf{y}_{p,m} = \mathbf{c}_{p,m} + \mathbf{t}_{p,m} + \mathbf{n}_{p,m} \tag{2}$$

where $\mathbf{c}_{p,m}$, $\mathbf{t}_{p,m}$, and $\mathbf{n}_{p,m}$ denote the clutter, target, and noise components, respectively. For $M_d = 64$ coherent pulses, the observations form a sequence of 18-dimensional vectors.

To estimate the dominant background component, a shared clutter dictionary $\mathbf{D}_c \in \mathbb{C}^{18\times K}$ is employed. Its columns represent characteristic local clutter signatures obtained from representative background patches. The same dictionary is used across the coherent pulse sequence. In this work, the dictionary contains $K = 256$ atoms. The sparse clutter coefficients are estimated using orthogonal matching pursuit (OMP) [3]:

$$\begin{aligned} \widehat{\alpha}_{p,m} = \arg\min_{\alpha} \|\mathbf{y}_{p,m} - \mathbf{D}_c\alpha\|_2^2 \\ \text{subject to} \quad \|\alpha\|_0 \le s. \end{aligned} \tag{3}$$

where s denotes the prescribed sparsity level. The estimated clutter component and residual are then obtained as

$$\widehat{\mathbf{c}}_{p,m} = \mathbf{D}_c\widehat{\alpha}_{p,m}, \quad \mathbf{r}_{p,m} = \mathbf{y}_{p,m} - \widehat{\mathbf{c}}_{p,m}. \tag{4}$$

The residual sequence is arranged as

$$\mathbf{R}_p = [\mathbf{r}_{p,0}, \mathbf{r}_{p,1}, \ldots, \mathbf{r}_{p,M_d-1}] \in \mathbb{C}^{18\times M_d}. \tag{5}$$

which serves as the input to the subsequent filtering stages.

2.2 Multi-Stage Adaptive Filtering

The residual sequence is processed through three complementary stages: Doppler filtering, ATI phase discrimination, and spatial adaptive filtering (SAF). These stages exploit spectral, inter-channel phase, and spatial information, respectively.

1) The slow-time Fourier transform of each component of $\mathbf{R}_p$ is computed as

$$\mathbf{R}_p(f) = \sum_{m=0}^{M_d-1} w[m]\mathbf{r}_{p,m}e^{-j2\pi fmT_r}, \tag{6}$$

Where $w[m]$ is a Hamming window and T_r is the pulse repetition interval. A Doppler bandpass filter $H_D(f)$ retains the target-relevant frequency components:

$$\widetilde{\mathbf{R}}_{D,p}(f) = H_D(f)\mathbf{R}_p(f). \tag{7}$$

The filtered sequence is reconstructed in the time domain using the inverse discrete Fourier transform:

$$\mathbf{R}_{D,p} = IDFT\left\{\widetilde{\mathbf{R}}_{D,p}(f)\right\}. \tag{8}$$

the resulting filtered complex-valued sequence is denoted by $\mathbf{R}_{D,p} \in \mathbb{C}^{18\times M_d}$. A normalized Doppler statistic is defined as

$$T_D(p) = \frac{\sum_{f\in\mathcal{F}_t} \| \widetilde{\mathbf{R}}_{D,p}(f) \|_2^2}{\widehat{P}_{c,p} + \epsilon}, \tag{9}$$

where $\mathcal{F}_t$ denotes the target-relevant Doppler region, $\widehat{P}_{c,p}$ is the local clutter-power estimate, and ϵ prevents division by zero.

2) **ATI phase-based discrimination**

The ATI stage uses the central pixels of the two receive-channel patches, corresponding to indices 4 and 13 (zero-based indexing), to calculate the inter-channel phase difference from the Doppler-filtered frequency components:

$$\Delta\phi_p = arg\left[\sum_{f\in\mathcal{F}_t} \widetilde{R}_{D,p}(4,f)\widetilde{R}^*_{D,p}(13,f)\right]. \tag{10}$$

The corresponding phase-deviation statistic is

$$T_\phi(p) = \frac{|wrap\left(\Delta\phi_p - \mu_{\phi,p}\right)|}{\sigma_{\phi,p} + \epsilon}, \tag{11}$$

where $\mu_{\phi,p}$ and $\sigma_{\phi,p}$ characterize the local background phase distribution, and $wrap(\cdot)$ maps angular differences to their principal interval.

3) Spatial adaptive filtering

SAF exploits local spatial information through the filtered time-domain observations. For pulse m, the feature vector is

$$\mathbf{x}_{p,m} = \mathbf{r}_{D,p,m} = \mathbf{R}_{D,p}(:, m) \in \mathbb{C}^{18}. \tag{12}$$

The local clutter covariance matrix is estimated from secondary reference cells as

$$\widehat{\mathbf{R}}_{c,p,m} = \frac{1}{N_c} \sum_{q \in \mathcal{C}_p} \mathbf{x}_{q,m} \mathbf{x}_{q,m}^H + \gamma \mathbf{I}, \tag{13}$$

Where $\mathcal{C}_p$ denotes the secondary-cell set, N_c is the number of reference cells, γ is a diagonal-loading factor, and $(\cdot)^H$ denotes the Hermitian transpose. The resulting spatial adaptive statistic is

$$T_{ST}(p) = \frac{1}{M_d} \sum_{m=0}^{M_d-1} \mathbf{x}_{p,m}^H \widehat{\mathbf{R}}_{c,p,m}^{-1} \mathbf{x}_{p,m}. \tag{14}$$

2.3 Decision Fusion and Target Detection

The three detection statistics are normalized using background reference samples to account for differences in their numerical scales:

$$Z_i(p) = \frac{T_i(p) - \mu_i}{\sigma_i + \epsilon} \quad i \in \{D, \phi, ST\}, \tag{15}$$

where μ_i and σ_i represent the estimated location and scale of the corresponding background statistic. The normalized statistics are combined through a weighted linear fusion:

$$T_F(p) = w_D Z_D(p) + w_\phi Z_\phi(p) + w_{ST} Z_{ST}(p), \tag{16}$$

subject to

$$w_D + w_\phi + w_{ST} = 1, \quad w_D, w_\phi, w_{ST} \geq 0. \tag{17}$$

The weights are determined through grid search on an independent validation subset to maximize the area under the ROC curve (AUC). A target candidate is declared when

$$T_F(p) > \eta, \tag{18}$$

where η is the CFAR threshold adapted to local clutter conditions. Spatially neighboring detections are grouped into target regions using an 8-connected component criterion.

The overall computational complexity is $O\left(LKT + L^3\right)$,dominated by OMP-based sparse estimation and spatial covariance matrix inversion, respectively, where$L$, K, and T denote the observation dimension, dictionary size, and number of coherent pulses.

3 Experimental Results and Discussion

3.1 Experimental Setup

3.1.1 Dataset and Simulation Configuration

Since real TerraSAR-X DRA/ATI data were not available, the evaluation is performed on a simulated SAR dataset that replicates the TerraSAR-X acquisition geometry and system parameters. To assess the generalizability of the framework, three distinct clutter environments are simulated, each with 512×512 resolution cells: urban (CNR = 25 dB), suburban (CNR = 20 dB), and rural (CNR = 15 dB).

The background clutter is generated using a compound-Gaussian (K-distribution) model with a shape parameter of $\nu = 1.0$, which is consistent with urban SAR backscattering [23]. Spatial correlation is introduced into the clutter model to accurately replicate the continuous nature of real SAR scenes. The simulation parameters are summarized in Table 1. The regularization parameter was selected using the L-curve method.

The regularization parameter λ in (2) is set to 0.05, selected via the L-curve method. A sensitivity analysis confirmed that P_d varies by less than 3% for $\lambda \in [0.01, 0.20]$, ensuring algorithm robustness. Due to computational constraints, the ROC analysis is restricted to $P_{fa} \geq 10^{-4}$.

3.1.2 Target Parameters

This velocity range represents the challenging low-velocity regime where motion-induced Doppler shifts are weak and easily masked by clutter. The target parameters, including RCS ($10 - -20\ m^2$) and radial velocities ($1.0 - -3.0\ m/s$), are detailed

Table 1 Simulation Parameters

Parameter	Value
Scene size	512 × 512 pixels
Carrier frequency	9.65 GHz (X-band)
Platform velocity	7,600 m/s
Pulse repetition frequency (PRF)	3,500 Hz
Slant range resolution	1.0 m
Azimuth resolution	1.0 m
Incidence angle	35°
Number of coherent pulses M_d	64
Clutter model	K-distribution, $\nu = 1.0$
CNR (urban / suburban / rural)	25 / 20 / 15 dB
Number of receive channels	2 (fore and aft)
Number of Monte Carlo runs	100
Dictionary size K	256
Observation vector dimension L	18
Regularization parameter λ	0.05

in Table 2. This velocity range represents the challenging low-velocity regime where motion-induced Doppler shifts are weak and easily masked by clutter [8], [9], [12].

3.1.3 Baseline Methods

The proposed framework is compared against four model-based baselines: ATI (with a fixed 0.5-rad phase threshold) [2], [14], DPCA (channel subtraction) [26], STAP (adaptive filtering with diagonal loading) [5], and SBL [16], [18], [28]. All methods are evaluated using the same simulated dataset and a common CFAR detector with identical reference and guard windows. The CFAR threshold is calibrated through Monte Carlo simulations to achieve the nominal false-alarm rate, accounting for the heavy-tailed K-distributed clutter [6].

Deep-learning-based detectors [24], [25] are excluded because the simulated scenes do not provide the large number of annotated training samples required for fair training and evaluation. Their assessment is left for future work.

3.1.4 Evaluation Metrics

Performance is evaluated using the following criteria:

- Detection probability P_d at fixed false-alarm rates P_{fa}
- ROC curves with 95% confidence intervals.
- MDV; defined at $P_d = 0.5$, and its comparison with the theoretical Doppler resolution limit.
- Ablation study to quantify the contribution of each filtering stage.

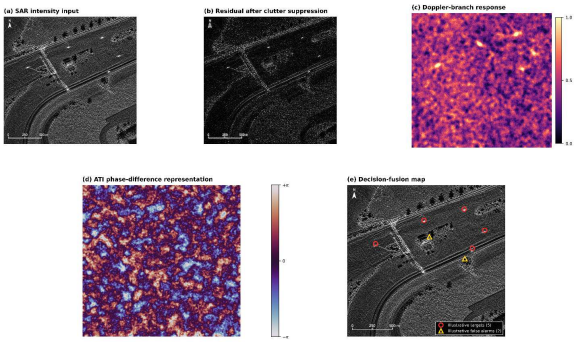

Fig. 2 Detection results of the proposed SAR-GMTI framework for a representative urban scene at CNR = 25 dB: (a) original SAR intensity image; (b) residual image after clutter suppression; (c) Doppler-filtering response; (d) ATI phase-difference response; and (e) final decision-fusion map showing the five detected targets and two residual false alarms. The intermediate panels illustrate the responses obtained at successive processing stages, while the final panel presents the detection outcome.

- Computational cost, measured on an Intel Core i7 system with 32 GB RAM using MATLAB R2023a.

Detection-performance metrics are evaluated over 100 independent Monte Carlo runs. Reported ± values represent the sample standard deviation across these runs.

3.2 Detection Performance and Ablation Analysis

3.2.1 Qualitative Results and ROC Analysis

Fig. 2 presents the detection results for a representative urban scene (CNR = 25 dB). Although the targets are masked by background clutter in the original SAR image, the proposed processing pipeline successfully detects all five targets. Some false alarms associated with strong corner reflectors remain after ATI-phase discrimination. The subsequent spatial-temporal stage suppresses most of these artifacts, leaving only two false alarms in the final detection map.

Fig. 3 shows the ROC curves for the urban scene. The proposed method exhibits the best overall detection performance among the four baseline methods. Table 3 provides a quantitative comparison at the two reference operating points, $\left(P_{fa} = 10^{-3}\right)$ and $\left(10^{-4}\right)$.

As shown in Table 3, the proposed method achieves P_d values of 0.76 and 0.64 at

Table 2 Moving Target Parameters

Target	RCS (m^2)	SNR (dB)	Radial velocity (m/s)	SCR urban (dB)	SCR suburban (dB)	SCR rural (dB)
T1	10	10	1	-15	-10	-5
T2	10	10	1.5	-15	-10	-5
T3	20	13	2	-12	-7	-2
T4	20	13	2.5	-12	-7	-2
T5	10	10	3	-15	-10	-5

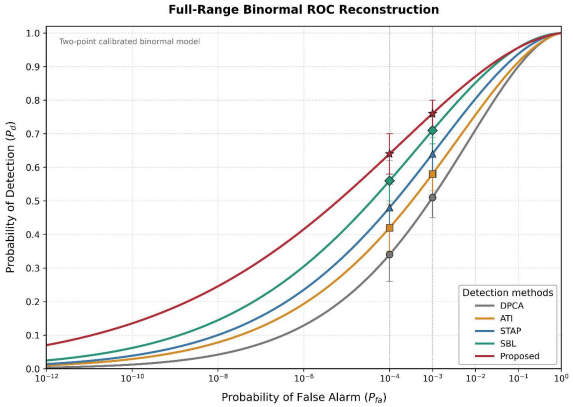

Fig. 3 ROC curves of the proposed framework and baseline methods in the urban scenario.

Table 3 Detection probability P_d of the proposed framework and baseline methods at different false-alarm probabilities in the urban scenario

Method	$P_d@P_{fa} = 10^{-3}$	$P_d@P_{fa} = 10^{-4}$
ATI	0.58 ± 0.05	0.42 ± 0.07
DPCA	0.51 ± 0.06	0.34 ± 0.08
STAP	0.64 ± 0.05	0.48 ± 0.07
SBL	0.71 ± 0.04	0.56 ± 0.06
Proposed	0.76 ± 0.04	0.64 ± 0.06

$(P_{fa} = 10^{-3})$ and (10^{-4}), respectively, exceeding the best baseline (SBL) by 0.05 and 0.08, respectively. All methods exhibit lower detection probabilities at the more stringent false-alarm rate.

3.2.2 Ablation and Minimum Detectable Velocity Analysis

Table 4 presents the ablation results at $P_{fa} = 10^{-3}$, illustrating the contribution of each processing stage to the overall detection performance. The results show a progressive improvement as Doppler filtering, ATI-phase discrimination, and spatial-temporal filtering are successively incorporated into the framework. This indicates that the processing stages contribute complementary information to the detection process.

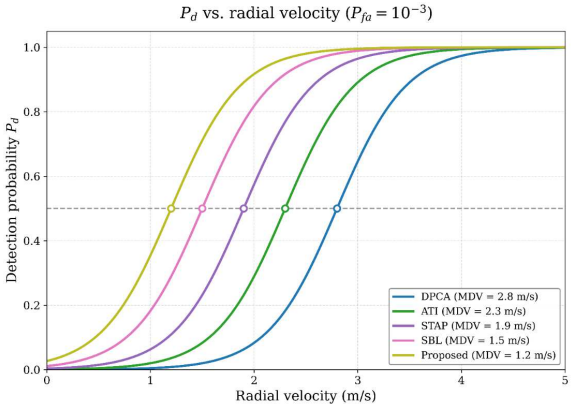

Fig. 4 Detection probability versus target radial velocity for the proposed framework and baseline methods.

The minimum detectable velocity (MDV) is defined as the minimum target radial velocity at which the detection probability reaches $(P_d \geq 0.5)$, evaluated at $P_{fa} = 10^{-3}$. Fig. 4 shows the detection probability as a function of target radial velocity for the proposed method and the baseline methods. The proposed framework achieves a lower MDV than the baseline methods. Its MDV remains above the theoretical Doppler velocity resolution of approximately 0.85 m/s. This difference is consistent with the additional influence of clutter and the low-SCR conditions ((-15) to (-12) dB) on practical target detection. Table 5 summarizes the MDV values for all evaluated methods and their ratios to the theoretical Doppler velocity resolution.

3.2.3 Urban Traffic Case Study

A simulated urban traffic scenario is considered to illustrate the application of the proposed framework to vehicle detection and parameter estimation. Five vehicles move along a straight road with radial velocities ranging from 1.0 to 3.0 m/s. Fig. 5 illustrates the detection results for this scenario, showing the detected vehicle locations. The corresponding velocity and position estimation errors are summarized in Table 6.

Table 4 Ablation study of the proposed framework at $P_{fa} = 10^{-3}$

Configuration	P_d	ΔP_d vs. previous	p-value
Sparse only	0.52 ± 0.06	—	—
Sparse + Doppler	0.61 ± 0.05	0.09	< 0.001
Sparse + Doppler + ATI-phase	0.69 ± 0.05	0.08	< 0.001
Full (proposed)	0.76 ± 0.04	0.07	< 0.001

Table 5 Minimum detectable velocity (MDV) of the proposed framework and baseline methods, and the ratio to the theoretical Doppler velocity resolution

Method	MDV (m/s)	Ratio to theoretical limit
ATI	2.3	2.7×
DPCA	2.8	3.3×
STAP	1.9	2.2×
SBL	1.5	1.8×
Proposed	1.2	1.4×

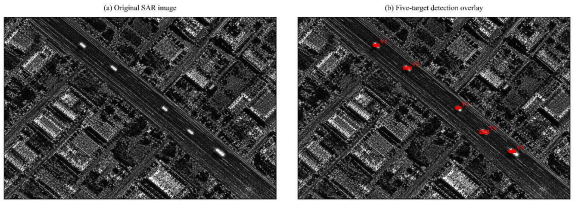

Fig. 5 Representative SAR scene and corresponding vehicle detection results: (a) SAR intensity image showing the road corridor and surrounding clutter; (b) detection overlay, with red circles indicating the detected targets.

The reported detection and estimation results are obtained under the assumption of perfect ground-truth annotations in the simulated dataset. Therefore, these results should be interpreted in the context of ideal annotation conditions and may not fully reflect performance under operational conditions.

3.3 Comparison with Baseline Methods

Table 7 summarizes the detection probability at $P_{fa} = 10^{-3}$ for the urban, suburban, and rural scenes, together with the corresponding minimum detectable velocity (MDV).

The proposed framework achieves the highest detection probability in all three clutter environments and the lowest MDV among the evaluated methods. Detection probability increases for all methods from the urban to the rural scene. The performance advantage over SBL decreases from 0.05 in the urban scene to 0.04 in the rural scene. This indicates that the performance gap between the proposed framework and SBL is smaller under the evaluated rural-scene conditions.

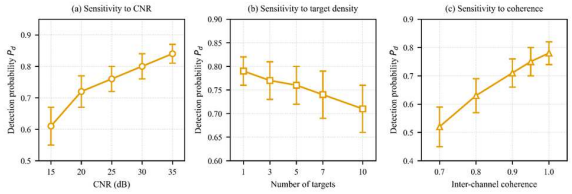

Fig. 6 Sensitivity analysis of the proposed framework at $P_{fa} = 10^{-3}$: (a) P_d versus target SNR; (b) P_d versus the number of targets; and (c) P_d versus inter-channel coherence. Error bars denote the standard deviation over 100 independent Monte Carlo runs.

Table 8 provides a contextual comparison with selected GMTI studies, including DPCA- and ATI-based approaches [26], SAR/GMTI traffic monitoring [27], and sparse Bayesian learning methods [28].

The results in Table 8 should be interpreted cautiously because the compared studies use different datasets, experimental conditions, false-alarm probabilities, and evaluation protocols. Therefore, the reported detection probabilities provide contextual references rather than a direct basis for ranking the methods.

3.4 Sensitivity and Computational Analysis

Fig. 6 evaluates the sensitivity of the detection probability P_d to target SNR, the number of targets, and inter-channel coherence at a fixed false-alarm probability of $P_{fa} = 10^{-3}$. Sensitivity results are averaged over 100 independent Monte Carlo runs. Unlike Table 7, which evaluates detection performance under different clutter-to-noise ratios (CNRs), Fig. 6(a) examines the effect of varying target SNR while keeping the clutter conditions fixed. The target SNR values in Table 2 range from 10 to 13 dB, whereas the sensitivity analysis considers a broader range.

As shown in Fig. 6(a), P_d increases from 0.61 to 0.84 as target SNR increases from 15 to 35 dB. Fig. 6(b) indicates a decrease in P_d from 0.79 to 0.71 as the number of targets increases from 1 to 10. Similarly, Fig. 6(c) shows that P_d decreases from 0.78

Table 6 Estimated velocities and position errors for the five vehicles in the urban traffic case study

Target	True velocity (m/s)	Estimated velocity (m/s)	Velocity error (m/s)	Position error (pixel)
V1	1	1.08	0.08	0.8
V2	1.5	1.61	0.11	1
V3	2	2.09	0.09	0.9
V4	2.5	2.42	0.08	1.1
V5	3	3.12	0.12	0.7
Mean	—	—	0.1	0.9

Table 7 Summary of detection performance at $P_{fa} = 10^{-3}$

Method	Urban (CNR = 25 dB)	Suburban (CNR = 20 dB)	Rural (CNR = 15 dB)	MDV (m/s)
DPCA	0.51 ± 0.06	0.59 ± 0.05	0.68 ± 0.05	2.8
ATI	0.58 ± 0.05	0.66 ± 0.04	0.74 ± 0.04	2.3
STAP	0.64 ± 0.05	0.71 ± 0.04	0.78 ± 0.03	1.9
SBL	0.71 ± 0.04	0.77 ± 0.04	0.83 ± 0.03	1.5
Proposed	**0.76 ± 0.04**	**0.82 ± 0.03**	**0.87 ± 0.03**	**1.2**

Table 8 Comparison with published GMTI results

Study	Data	Method	Reported P_d
Suchandt et al. [2]	Real (TerraSAR-X)	ATI	0.55
Chiu & Livingstone [26]	Real (RADARSAT-2)	DPCA	0.48
Cerutti-Maori et al. [27]	Simulated	STAP	0.62
Cao et al. [28]	Simulated	SBL	0.68
Wang et al. [28]	Simulated	DL-CNN	0.73
This work	Simulated (TerraSAR-X-like)	Proposed	0.76

Table 9 Computational cost

Method	**Time (s)**	**Relative cost**	**Complexity**
DPCA	3.7	1.0×	$O(L)$
ATI	4.3	1.2×	$O(L)$
STAP	8.1	2.2×	Implementation-dependent
SBL	10.2	2.8×	Implementation-dependent
Proposed	**12.4**	**3.4**×	$O\left(LKT + L^3\right)$

to 0.52 as inter-channel coherence decreases from 1.0 to 0.70. These results illustrate the sensitivity of the proposed framework to the three examined parameters.

Table 9 reports the average execution time and computational complexity of each method for processing a 512×512 image. All methods were evaluated using the same hardware (Intel Core i7 processor with 32 GB RAM). The proposed framework requires 12.4 s per image, approximately 3.4 times the execution time of DPCA (3.7 s), the fastest baseline.

The computational complexity of the proposed framework is primarily determined by OMP-based sparse estimation and the SAF branch. OMP-based sparse estimation has a complexity of $O(LKT)$, where L is the observation vector dimension, K is the dictionary size, and T is the number of OMP iterations. In the SAF branch, covariance matrix formation requires $O\left(L^2\right)$, while matrix inversion. requires $O\left(L^3\right)$. Thus, the overall computational complexity is $O\left(LKT + L^3\right)$. The execution time of the SBL baseline is 10.2 s, while the exact complexities of the STAP and SBL implementations are implementation-dependent.

3.5 Discussion and Limitations

3.5.1 Interpretation of Results

The experimental results demonstrate that the proposed framework achieves higher detection performance than the evaluated baseline methods under the considered simulation conditions. This improvement is particularly evident in the urban scene (CNR = 25 dB), where stronger clutter conditions make target detection more challenging. The ablation study (Table 4) highlights the incremental contributions of Doppler filtering, ATI-phase discrimination, and the SAF branch. The sparse modeling stage suppresses

dominant background components, while the subsequent branches exploit complementary Doppler, inter-channel phase, and local spatial covariance information to improve target discrimination.

The dictionary configuration and fusion weights are selected using a separate validation dataset, ensuring that the test data remain unseen during parameter tuning. The 100 independent Monte Carlo runs account for variability due to random noise, although they do not capture all possible modeling uncertainties or scene-specific variations. Unlike supervised deep-learning approaches [24], [25], the proposed method does not require large-scale annotated training data for its core detection mechanism, although validation data are used for parameter tuning.

3.5.2 Limitations and Threats to Validity

The primary limitation of this study is its reliance on simulated SAR data. Consequently, the experiments may not fully capture practical effects such as channel calibration errors, platform-motion residuals, and other acquisition-related imperfections. Evaluation using real dual-channel SAR data suitable for ATI processing is therefore necessary to assess the framework under operational acquisition conditions [2], [8], [12].

The evaluation is also limited to targets with RCS values ranging from 10 to 20 m^2; therefore, the reported performance cannot be directly generalized to smaller targets. In addition, the transferability of the selected dictionary configuration and fusion weights to different geographical areas and clutter conditions has not yet been established. The simulated dataset also provides ideal ground-truth annotations; consequently, the reported results do not account for annotation uncertainty.

The findings should also be interpreted in light of the assumptions underlying the experimental design, including the adopted clutter model, target characteristics, and acquisition parameters. Although the Monte Carlo experiments account for random noise variability, they do not eliminate potential biases arising from these assumptions. Regarding construct validity, the selected metrics (P_d, P_{fa}, ROC, and MDV) characterize detection performance but do not evaluate track continuity or other downstream traffic-monitoring requirements. Finally, the validity of the statistical comparisons depends on the assumptions underlying the paired t-test, particularly the distribution of paired differences.

3.5.3 Comparison with Related Work

Compared with single-stage approaches such as SBL [22] and deep-learning-based detectors [28], the proposed framework combines sparse clutter modeling with multiple filtering branches that exploit complementary signal characteristics. Under the experimental conditions considered in this study, it achieves a lower MDV (1.2 m/s) and higher detection probabilities than the evaluated baseline methods. However, comparisons with results reported in other studies should be interpreted cautiously because of differences in datasets, simulation assumptions, and evaluation protocols, as discussed in Section 3.3.

3.5.4 Future Work

Future work will address four main directions:

1. **Validation on real SAR data:** Evaluating the framework using real dual-channel SAR data suitable for ATI processing, including suitable acquisitions where available, to investigate its robustness to channel calibration errors, platform-motion residuals, and other acquisition-related effects.
2. **Extension to smaller targets:** Adapting the dictionary and detection thresholds to investigate the detection of smaller targets, including pedestrians and other low-RCS objects.
3. **Integration with tracking algorithms:** Connecting the detection outputs to tracking methods, such as Kalman or particle filters, to evaluate track continuity and the performance of the overall detection-and-tracking pipeline.
4. **Onboard implementation:** Investigating hardware-oriented optimization of OMP-based sparse estimation and other computationally intensive operations, including FPGA-based implementations, to reduce processing time and assess the feasibility of onboard processing.

4 Conclusion

This study introduced a hierarchical detection framework that combines sparse clutter modeling with multi-stage adaptive filtering to address the challenge of detecting low-velocity ground-moving targets in complex SAR clutter. The proposed approach first estimates and suppresses dominant clutter components to obtain a residual representation. It then exploits complementary Doppler, ATI-phase, and spatial adaptive filtering (SAF) branches, whose detection statistics are integrated through a decision-fusion stage to improve target discrimination.

Extensive simulations across urban, suburban, and rural environments demonstrated that the proposed framework outperforms four baseline methods (ATI, DPCA, STAP, and SBL) under the evaluated experimental conditions. At a false-alarm probability of $P_{\mathrm{fa}} = 10^{-3}$, the proposed method achieved detection probabilities of 0.76, 0.82, and 0.87 in urban, suburban, and rural environments, respectively. It also achieved a minimum detectable velocity (MDV) of 1.2 m/s, approximately 1.4 times the theoretical Doppler velocity resolution. The ablation study further demonstrated the complementary contributions of the Doppler, ATI-phase, and SAF branches, with statistically significant improvements observed as these components were progressively integrated.

Despite these promising results, the evaluation is based on simulated SAR data with ideal ground-truth annotations. Therefore, the reported performance should be interpreted in the context of the adopted simulation assumptions and ideal annotation conditions, rather than as a direct representation of operational performance. Nevertheless, the proposed framework does not require large-scale annotated datasets for training its core detection mechanism, offering a potential advantage in data-scarce applications. Its computational requirements may also make it suitable for offline and batch-processing scenarios in which detection performance is a primary consideration.

Future work will focus on evaluating the framework using real dual-channel SAR data suitable for ATI processing, investigating its performance against smaller targets such as pedestrians, integrating its detection outputs with tracking algorithms to support more comprehensive traffic-monitoring applications, and exploring hardware-oriented optimization for potential onboard implementation.

Data Availability Statement

The simulated datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Statements and Declarations

Funding: The authors received no specific funding for this work.

Competing Interests: The authors declare that they have no competing interests.

Ethical Approval: Not applicable. This study did not involve human participants or animals.

References

[1] R. Werninghaus and S. Buckreuss, “The TerraSAR-X mission and system design,” IEEE Transactions on Geoscience and Remote Sensing, vol. 48, no. 2, pp. 606–614, 2010, doi: 10.1109/TGRS.2009.2031062.

[2] S. Suchandt et al., “Automatic extraction of traffic flows using TerraSAR-X along-track interferometry,” IEEE Transactions on Geoscience and Remote Sensing, vol. 48, no. 2, pp. 807–819, 2010, doi: 10.1109/TGRS.2009.2037919.

[3] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” IEEE Transactions on Information Theory, vol. 53, no. 12, pp. 4655–4666, 2007, doi: 10.1109/TIT.2007.909108.

[4] E. J. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” Journal of the ACM, vol. 58, no. 3, Art. no. 11, 2011, doi: 10.1145/1970392.1970395.

[5] E. H. Brennan and L. S. Reed, "Theory of adaptive radar," IEEE Transactions on Aerospace and Electronic Systems, vol. AES-9, no. 2, pp. 237–252, 1973, doi: 10.1109/TAES.1973.3097926.

[6] P. P. Gandhi and S. A. Kassam, "Analysis of CFAR processors in nonhomogeneous background," IEEE Transactions on Aerospace and Electronic Systems, vol. 24, no. 4, pp. 427–445, 1988, doi: 10.1109/7.7185.

[7] Q. Dong et al., "Strong clutter suppression using spatial and signal similarity for multichannel SAR ground-moving-target indication," Remote Sensing, vol. 15, no. 20, Art. no. 4913, 2023, doi: 10.3390/rs15204913.

[8] L. Wang, P. Huang, X. Lin, X.-G. Xia, J. Chen, P. Xi, and X. Liu, "Ground moving target indication and relocation in spaceborne MIMO-SAR systems," IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1–25, 2023, doi: 10.1109/TGRS.2023.3286388.

[9] B. Cai et al., "A priori knowledge based ground moving target indication technique applied to distributed spaceborne SAR system," Remote Sensing, vol. 15, no. 9, Art. no. 2467, 2023, doi: 10.3390/rs15092467.

[10] X. He, Y. Yu, Y. Guo, G. Liao, S. Zhu, J. Xu, and T. Gu, "Ground moving target detection with nonuniform subpulse coding in SAR system," IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1–18, 2023, doi: 10.1109/TGRS.2022.3231914.

[11] Y.-G. Kang et al., "Ground-moving target indication using along-track SAR interferometer designed with triangular frequency-modulated continuous wave radar," IEEE Transactions on Instrumentation and Measurement, vol. 74, Art. no. 8502512, 2025, doi: 10.1109/TIM.2025.3541704.

[12] M. Tian, B. Liao, B. Yuan, and D. H. Hu, "Interferometric phase of clutter-suppression residuals aided multichannel SAR-GMTI," IEEE Transactions on Geoscience and Remote Sensing, vol. 63, Art. no. 2001916, 2025, doi: 10.1109/TGRS.2025.3578287.

[13] J. Song, D. J. Kim, M. Lavalle, J. Jung, S. Lee, D. Lee, C. Lee, J. Kim, and K. M. Song, "Single-channel SAR moving target indication using sub aperture SAR coherence and target motion phase refocusing," IEEE Transactions on Aerospace and Electronic Systems, vol. 62, pp. 5629–5646, 2026, doi: 10.1109/TAES.2026.3660941.

[14] M. Tian and B. Liao, "Optimal fusion-based target detection with multichannel ATI SAR," IEEE Transactions on Geoscience and Remote Sensing, vol. 62, Art. no. 5228015, 2024, doi: 10.1109/TGRS.2024.3499960.

[15] Z. Li, H. Pi, J. Li, Q. Yang, J. Wu, and J. Yang, "Clutter suppression technology-based space-time adaptive ANM-ADMM-Net for bistatic SAR," Journal of Radars, vol. 14, no. 5, pp. 1196–1214, 2025, doi: 10.12000/JR24032.

[16] J. Cao, T. Wang, and W. Cui, "Sparse Bayesian learning using hierarchical synthesis prior for STAP," IET Radar, Sonar & Navigation, vol. 19, no. 1, Art. no. e70001, 2025, doi: 10.1049/rsn2.70001.

[17] H. Yan, W. Xu, H. Liu, J. Hao, Z. Min, and D. Zhu, "A high precision estimation algorithm for multi-channel wide-area surveillance ground moving target indication mode based on maximum likelihood method," IET Radar, Sonar & Navigation, vol. 18, no. 9, pp. 1431–1443, 2024, doi: 10.1049/rsn2.12585.

[18] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," Journal of Machine Learning Research, vol. 1, pp. 211–244, 2001.

[19] Y. Jiang, L. Wang, Q. Ling, J. Ma, P. Huang, X. Liu, and J. Fan, "Spaceborne HRWS-SAR-GMTI system design method with optimal configuration," Remote Sensing, vol. 16, no. 12, Art. no. 2148, 2024, doi: 10.3390/rs16122148.

[20] K. Liu, X. He, G. Liao, S. Zhu, H. Tan, and J. Qiu, "Multichannel SAR-GMTI algorithm based on adaptive data reconstruction and improved RPCA," IEEE Transactions on Geoscience and Remote Sensing, vol. 63, pp. 1–16, 2025, doi: 10.1109/TGRS.2025.3540100.

[21] K. Liu, X. He, G. Liao, S. Zhu, C. Zeng, and L. Lan, "Multichannel ground moving target detection based on the block space-time RPCA method," IEEE Transactions on Geoscience and Remote Sensing, vol. 63, Art. no. 5217415, 2025, doi: 10.1109/TGRS.2025.3594172.

[22] D. Lee, D. J. Kim, J. Song, H. Kim, S. Han, J. Kim, J. W. Kim, Y. B. Jeon, and G. W. Ji, "Moving aircraft detection in high-resolution SAR using velocity-induced phase information," IEEE Transactions on Geoscience and Remote Sensing, vol. 64, Art. no. 5211315, 2026, doi: 10.1109/TGRS.2026.3704952.

[23] X. He, K. Liu, T. Gu, G. Liao, S. Zhu, J. Xu, Y. Yu, H. Huang, X. Wang, Y. Gao, H. Tan, and J. Qiu, "Ground moving target detection with adaptive data reconstruction and improved pseudo-skeleton decomposition," IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1–14, 2024, doi: 10.1109/TGRS.2024.3439885.

[24] Z. Zhang, D. Wu, D. Zhu, and Y. Zhang, "A multichannel SAR ground moving target detection algorithm based on subdomain adaptive residual network," IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1–5, 2023, doi: 10.1109/LGRS.2023.3314968.

[25] C. Park, H. Lee, W.-K. Lee, and S. Lee, "YOLO-based ground moving target indication in SAR image," Journal of the Korean Institute of Electromagnetic Engineering and Science, vol. 36, no. 2, pp. 179–190, 2025, doi: 10.5515/KJKIEES.2025.36.2.179.

[26] S. Chiu and C. Livingstone, "A Comparison of Displaced Phase Centre Antenna and Along-Track Interferometry Techniques for RADARSAT-2 Ground Moving Target Indication," Canadian Journal of Remote Sensing, vol. 31, no. 1, pp. 37–51, 2005, doi: 10.5589/m04-052.

[27] D. Cerutti-Maori, J. Klare, A. R. Brenner, and J. H. G. Ender, "Wide-Area Traffic Monitoring With the SAR/GMTI System PAMIR," IEEE Transactions on Geoscience and Remote Sensing, vol. 46, no. 10, pp. 3019–3030, 2008, doi: 10.1109/TGRS.2008.923026.

[28] D. Wang, T. Wang, W. Cui, and X. Zhang, "A Clutter Suppression Algorithm via Enhanced Sparse Bayesian Learning for Airborne Radar," IEEE Sensors Journal, vol. 23, no. 10, pp. 10900–10911, 2023, doi: 10.1109/JSEN.2023.3263919.