Astronomical Techniques and Instruments, 2026

Article Open Access

SLR video weak target detection based on spatio-temporal variance fusion

Jinlong Hu[1], Yi Zhang[1], Tong Zou[1,2,3*]

[1]*Institute of Seismology*, *China Earthquake Administration*, *Wuhan* 430071, *China*

[2]*Wuhan Gravitation and Solid Earth Tides*, *National Observation and Research Station*, *Wuhan* 430071, *China*

[3]*Key Laboratory of Earthquake Geodesy*, *China Earthquake Administration*, *Wuhan* 430071, *China*

*Correspondence: tong.zou@163.com

Received: February 5, 2026; Accepted: March 24, 2026; Published Online: March 25, 2026; https://doi.org/10.3724/ati2026010; https://cstr.cn/32083.14.ati2026010

© 2026 Editorial Office of Astronomical Techniques and Instruments, Yunnan Observatories, Chinese Academy of Sciences. This is an open access article under the CC BY-NC-ND 4.0 license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Citation: Hu, J. L., Zhang, Y., Zou, T. 2026. SLR video weak target detection based on spatio-temporal variance fusion. *Astronomical Techniques and Instruments*, https://doi.org/10.3724/ati2026010.

Abstract: To address the critical detection challenges in Satellite Laser Ranging (SLR) observations—including random noise, fixed-pattern interference (e.g., hot pixels), dynamic trajectory confusion, strong signal masking, and low grayscale contrast—this paper proposes a weak target detection algorithm for SLR videos based on spatio-temporal variance fusion. First, a comprehensive preprocessing module is constructed, incorporating bad pixel mask restoration and column-wise statistical differencing to concurrently eliminate intrinsic sensor hot pixels and vertical stripe noise; meanwhile, a non-linear square mapping is employed to enhance the grayscale contrast of weak targets. Second, a dual-channel detection mechanism is designed by leveraging the complementary characteristics of spatio-temporal variance: the spatial variance channel utilizes temporal mean and spatial variance to jointly accumulate weak signals, addressing the invisibility issue caused by single-frame noise; the temporal variance channel utilizes the temporal fluctuation features of pixel intensity to physically distinguish moving targets from static high-brightness interference, thereby resolving the indistinguishability problem. Finally, the strong signal masking effect is effectively overcome through the fusion of spatio-temporal variance features and a dynamic saturation mechanism based on the noise floor. Extensive tests on 92 real observation videos involving 11 types of satellites demonstrate that the proposed algorithm achieves a high detection rate of 95.6% with a false alarm rate of only 3.3%. The experimental results indicate that the algorithm can effectively mitigate the effects of strong interference and low contrast, achieving high-precision localization and extremely low false-alarm detection for weak SLR targets. A demonstration video of the proposed method is available at https://doi.org/10.57760/sciencedb.j00149.00007.

Keywords: Satellite laser ranging; Image processing; Signal-to-noise ratio; Target detection

1. INTRODUCTION

Satellite Laser Ranging (SLR), by virtue of its extremely high monochromaticity and directionality, has become the most precise technical means for spatial target ranging and is widely applied in precision orbit determination and geodynamics research[1–3]. In the SLR operational workflow, video data is primarily used for auxiliary confirmation and closed-loop tracking support after the target enters the field of view (FOV). In practical observations, however, the signal-to-noise ratio (SNR) of SLR videos is relatively low. The imaging characteristics present several critical interference factors, many of which are associated with common hardware limitations. Random noise, including readout noise and photon noise, can easily submerge the relatively faint target, making it nearly invisible in a single frame. Sensor defects and fixed-pattern noise also introduce significant interference into the imaging process. For example, hot pixels appear as isolated bright spots with fixed positions and nearly constant brightness due to sensor imperfections. In addition, vertical stripe noise, namely fixed-pattern noise (FPN) distributed along the readout direction, may arise from non-uniformity in the readout circuitry of the charge-coupled device (CCD). Dynamic trajectory confusion may further occur because accidental streaks generated by non-target objects, as well as stars elongated by relative telescope motion, can be mistakenly identified as target trajectories. Moreover, the presence of active laser beams or laser scattering within the FOV may produce a strong signal masking effect, often referred to as the “darkness under the

lamp" phenomenon, causing weak satellite signals to be suppressed after global normalization. Finally, the target usually exhibits extremely low image contrast, with its signal amplitude only slightly higher than the background noise level, which further increases the difficulty of reliable detection.

These interference factors make it difficult for observers to manually identify target satellites. Utilizing image recognition technology for automated target recognition can effectively improve observation efficiency, reduce over-reliance on manual experience, and lay the foundation for achieving high-precision, unmanned automated SLR observations.

To address the aforementioned challenges, existing research is primarily divided into two categories: single-frame spatial-feature-based methods and multi-frame temporal-analysis-based methods. However, both exhibit limitations when faced with the unique imaging environment of SLR.

In terms of single-frame spatial features, early research often employed mathematical morphological filtering (e.g., Top-Hat transform[4]), utilizing specific structural elements to suppress the background. Subsequently, local contrast methods inspired by the human visual system (e.g., LCM[5], MPCM[6]) and models combined with low-rank sparse recovery theory (IPI[7]) have gradually become mainstream. These methods extract targets by enhancing local differences or separating low-rank and sparse matrices. However, these methods primarily rely on brightness and local texture features. In SLR observation data, grayscale differences of weak targets are often submerged, and it is difficult for these algorithms to physically distinguish true targets from noise pixels, leading to extremely high false alarm rates under strong noise backgrounds.

In the aspect of multi-frame temporal analysis, utilizing the temporal correlation of sequences is key to improving the SNR. Classic background modeling (e.g., GMM[8]) and optical flow methods are highly prone to misidentifying random noise as foreground under extremely low SNR conditions. Recent studies have analyzed the performance of satellite tracking algorithms specifically tailored for low SNR environments[9] and explored the capabilities of optical Commercial Off-The-Shelf (COTS) stations for wide-field detection of space objects[10]. To further solve the detection challenge, algorithms based on the "Track-Before-Detect (TBD)" framework have been proposed, such as Particle Filter[11] and 3D Matched Filter[12]. Although they can significantly improve the detection upper bound through energy accumulation, they are often accompanied by exponentially growing search space and computational cost, making it difficult to meet engineering real-time requirements. Furthermore, simple temporal linear stacking methods[13,14], while able to smooth random noise, fall into the "denoising vs. anti-interference contradiction"—while enhancing the target, they simultaneously enhance fixed-position hot pixels and bad points, failing to effectively eliminate FPN interference.

Although deep learning methods (e.g., You Only Look Once (YOLO)) have achieved remarkable success in general object detection tasks, traditional spatio-temporal statistical methods remain highly preferable for SLR video processing. Unlike data-driven models, these "white-box" approaches offer clear physical interpretability, simple implementation, and independence from large-scale annotated datasets. Furthermore, their low computational demands are well-suited for real-time, high-frame-rate processing on low-cost systems typical of SLR stations. Crucially, SLR targets are low-SNR, point-like objects lacking the stable geometric and textural features that deep learning heavily relies upon, giving traditional statistical methods an inherent advantage in this scenario.

This paper proposes an SLR video target recognition algorithm based on spatio-temporal variance fusion, which can effectively remove hardware and environmental interference, and significantly improve the detection rate.

2. ALGORITHM

The overall framework of the proposed SLR weak target detection algorithm is illustrated in Fig. 1. Considering the spatio-temporal characteristics of SLR image sequences, the algorithmic pipeline is designed as three cascaded modules: an image preprocessing module, a spatio-temporal variance feature extraction module, and a signal fusion and adaptive mapping module. Specifically, the input to the algorithm is a sequence of continuous SLR video frames. First, in the image preprocessing stage, a pre-calibrated mask generated from multi-frame averaging and binary threshold segmentation is employed to accurately locate and repair intrinsic bad pixels of the sensor through neighborhood interpolation. At the same time, column-wise statistical differencing is performed to suppress vertical stripe noise. A nonlinear square mapping is then introduced to stretch the image grayscale and enhance the contrast between weak signals and the background. Subsequently, the preprocessed data stream is fed into a dual-channel parallel architecture for spatio-temporal variance feature extraction. The spatial variance channel focuses on signal intensity accumulation, thereby highlighting weak target responses and suppressing interference from a statistical perspective, whereas the temporal variance channel emphasizes the fluctuation characteristics of signal intensity over time, enabling effective discrimination between hot-pixel interference and real moving targets according to their temporal behavior. Finally, the feature maps produced by the two channels are subjected to a dual-gate screening and fusion process to preserve high-confidence target signals. On this basis, an adaptive mapping mechanism guided by the noise floor is introduced to dynamically adjust the gain while suppressing background noise, thereby alleviating the strong signal masking effect and improving detection accuracy.

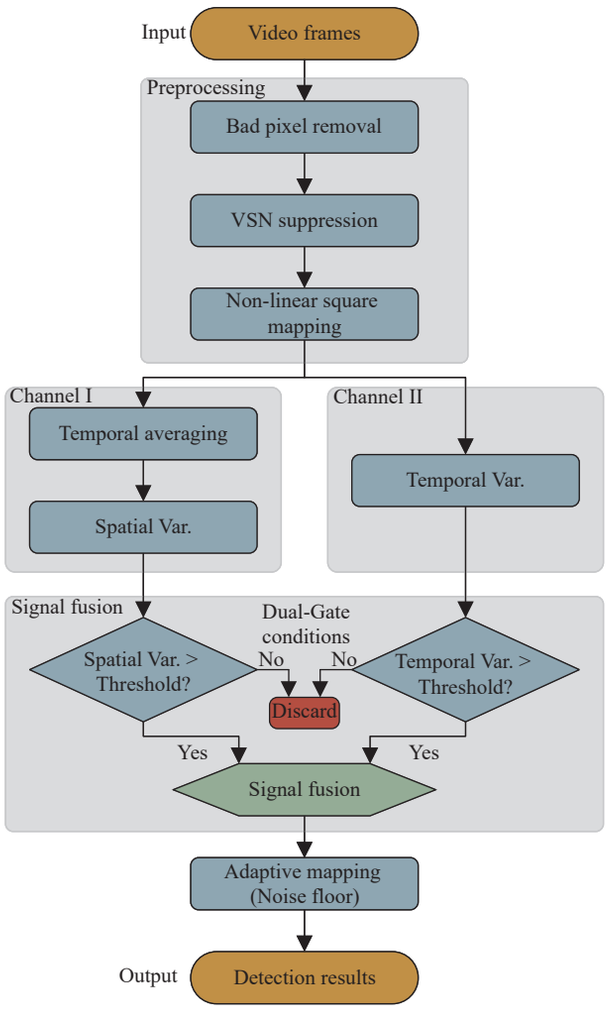


Fig. 1. Schematic diagram of the proposed algorithm architecture. VSN: Vertical Stripes Noise; Var.: Variance.

2.1. Vertical Stripe Noise Suppression

SLR utilizes CCD cameras, where the non-uniformity of the CCD readout circuits generates FPN. In the captured images, this noise primarily manifests as vertical stripe noise[15]. While this stripe noise is not prominent in a single frame, it becomes clearly visible after multi-frame stacking (as shown Fig. 2A), which significantly hinders subsequent algorithmic processing.

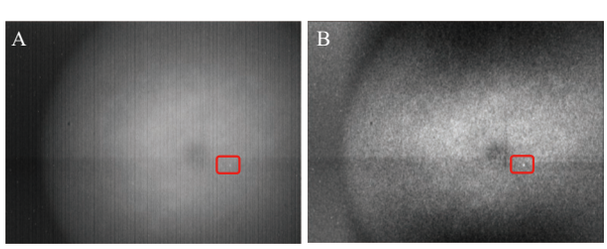


Fig. 2. Performance of stripe noise suppression. (A) Original image. (B) Denoised image.

It should be noted that the preprocessing steps described here, such as bad pixel repair and vertical stripe suppression, are not equivalent to standard dark-field or flat-field radiometric calibrations. While standard dark/flat calibrations compensate for fixed bias and sensor non-uniformity at the imaging system level, our algorithm-level compensation specifically targets residual sensor defects and fixed-pattern noise embedded within the observation videos to facilitate the subsequent weak target statistical detection.

In this paper, a column-wise statistical differencing method is adopted for correction. The algorithm calculates the grayscale mean M_j for each column j of a single frame and the global grayscale mean M_{global} of the entire image. The offset of each column relative to the global background is then subtracted from the original image I:

$$I'(i,j) = I(i,j) - (M_j - M_{\text{global}}) , \tag{1}$$

where $I(i,j)$ represents the grayscale value of the pixel at row i and column j in the original image, M_j represents the mean grayscale value of the j-th column, M_{global} represents the global mean grayscale value of the entire frame, and $I'(i,j)$ represents the corrected pixel value.

This step effectively eliminates the masking effect of stripe noise on weak signals while maintaining the baseline background brightness of the image. The suppression results are illustrated in Fig. 2.

2.2. Nonlinear Square Mapping

To address the challenge of extremely low brightness and contrast inherent to satellite targets, this study adopts a strategy combining "mean smoothing + square mapping" to construct the instantaneous energy map $S(x,y)$:

$$S(x,y) = [Smooth_{3\times3}(I_{\text{clean}})]^2. \tag{2}$$

The squaring operation transforms minute linear grayscale differences into significant quadratic non-linear differences. For normalized grayscale values, the grayscale gradients in high-brightness regions (potential targets) are significantly stretched, whereas low-grayscale background noise regions are further compressed. From a mathematical perspective, the contrast gain is directly proportional to the original grayscale intensity I (since $\mathrm{d}I^2/\mathrm{d}I = 2I$). This implies that the brighter the target, the more significant the contrast enhancement it receives.

Compared to higher-order transformations such as cubic mapping (I^3), the selection of square mapping in this paper is primarily based on two main considerations, namely signal preservation and computational efficiency. On the one hand, higher-order mappings tend to suppress low-grayscale regions too aggressively, which may over-compress weak targets at critical grayscale levels and lead to signal loss. On the other hand, the squaring operation imposes a much lower computational burden than higher-order power operations, making it more suitable for the real-time processing requirements of high-frame-rate SLR video streams.

This strategy effectively suppresses background noise while preliminarily highlighting the contour of weak targets. The contrast enhancement effect is illustrated in Fig. 3.

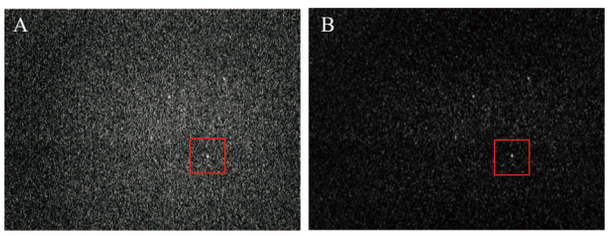


Fig. 3. Comparison of nonlinear square mapping effects. The weak target is indicated by the red box. (A) Input image. (B) Processed image.

2.3. Spatio-temporal Variance Feature Extraction

This is the core component of the proposed algorithm. To address the complex scenarios in SLR images where targets coexist with various interferences (e.g., random noise and hot pixels), this paper leverages the differences in spatio-temporal statistical characteristics between targets and interferences. Based on a sliding time window of length $T = 100$ frames, a dual-channel parallel detection architecture incorporating both spatial variance and temporal variance is constructed.

2.3.1. *Spatial variance calculation*

The spatial variance calculation primarily addresses two major challenges: weak signal intensity and dynamic trajectory confusion. A "temporal averaging followed by spatial variance" approach is designed as follows:

(1) Suppression of dynamic trajectories: While the target position remains relatively stationary in the video, stars and non-target objects appear as dynamic trajectories, which hinders target identification. By applying temporal averaging, these dynamic trajectories can be effectively suppressed while the steady-state satellite signal is preserved, thereby resolving the issue of dynamic trajectory confusion.

(2) Enhancement of weak signals: Following the suppression of random noise and dynamic interference via temporal averaging, the target region exhibits a "high-center, low-periphery" energy distribution. To exploit this, the local spatial variance within a neighborhood is calculated. Specifically, a 7×7 sliding window is employed to compute this local variance.

Mathematically, the spatial variance $S_{\text{space}}(x,y)$ is defined as:

$$S_{\text{space}}(x,y) = \frac{1}{N}\sum_{(i,j)\in W_{x,y}}\left[I_{\text{avg}}(i,j) - \mu_{\text{local}}(x,y)\right]^2, \quad (3)$$

where $W_{x,y}$ denotes the 7×7 local sliding window centered at (x,y), $N = 49$ is the total number of pixels within the window, and $\mu_{\text{local}}(x,y)$ represents the mean value of I_{avg} within this local window. This step utilizes statistical differences to further amplify the contrast between the steady-state target signal and the flat background, effectively addressing the challenge of weak signal detection.

2.3.2. *Temporal variance calculation*

There are distinct differences between hot pixels and actual targets in CCD observations. Hot pixels and bad pixels remain at fixed positions and usually exhibit nearly constant brightness over time; therefore, their temporal variance approaches zero. In contrast, actual targets undergo slow displacement across the FOV and are often accompanied by brightness scintillation, so their pixel values fluctuate significantly along the time axis, resulting in relatively high temporal-variance responses (> 0).

Calculation method: Traditional temporal accumulation tends to simultaneously enhance fixed-position hot pixels, leading to false alarms. To mitigate this, this channel utilizes temporal variance to measure the degree of pixel fluctuation within the sliding window, defined mathematically as $S_{\text{time}}(x,y)$:

$$S_{\text{time}}(x,y) = \sigma_t^2(x,y) = \frac{1}{T}\sum_{k=1}^{T}\left[I_k(x,y) - I_{\text{avg}}(x,y)\right]^2, \quad (4)$$

where $I_k(x,y)$ denotes the original grayscale value at pixel coordinates (x,y) in the k-th frame of the input video sequence, and $I_{\text{avg}}(x,y)$ denotes the mean grayscale value of the consecutive image frames at the same coordinates (x,y).

Through this channel, the algorithm can accurately identify and eliminate "pseudo-targets"—pixels that appear bright in the spatial variance map but lack temporal fluctuations. As shown in Fig. 4, the temporal variance effectively identifies pseudo-target points present in the spatial variance map.

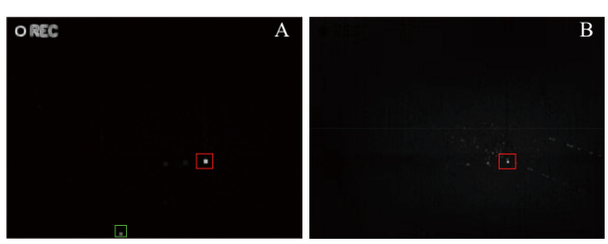


Fig. 4. Performance comparison between spatial and temporal variance. (A) Spatial variance. (B) Temporal variance, the input image is Fig. 3B).

In Fig. 4A, the green box indicates thermal noise (hot pixels), which is successfully suppressed in the temporal variance result on the right. Conversely, in the right panel, dense small points visible near the target represent dynamic noise, which is effectively suppressed in the spatial variance result on the left.

2.4. Signal Fusion

As demonstrated in the previous sections, the spatial and temporal channels effectively suppress different types

of interference. To further eliminate residual noise and highlight the target, this study constructs a feature fusion mechanism based on dual-gate verification. The specific signal fusion mask is defined as follows:

$$Mask(x,y)=\begin{cases}1 & \text{if } S_{\text{space}}>\mu_s+5\sigma_s \\ & \wedge\, S_{\text{time}}>\mu_d+3\sigma_d \\ 0 & \text{otherwise}\end{cases}, \tag{5}$$

where S_{space} and S_{time} denote the spatial and temporal variance feature maps generated in Section 2.3, respectively; μ_s and σ_s denote the mean and standard deviation of the background region in the spatial feature map; and μ_d and σ_d denote the mean and standard deviation of the background region in the temporal feature map.

Based on this mask, the intensity of the preserved target signal, denoted as $S_{\text{fusion}}(x,y)$, is defined mathematically as:

$$S_{\text{fusion}}(x,y)=S_{\text{space}}(x,y)\cdot Mask(x,y). \tag{6}$$

This gating mechanism incorporates physical constraints at two levels. First, the spatial intensity constraint requires S_{space} to be significantly higher than the background noise floor, with the threshold set to $\mu_s+5\sigma_s$, so as to suppress random noise interference within individual frames. Second, the temporal dynamic constraint requires S_{time} to exhibit significant temporal volatility, with the threshold set to $\mu_d+3\sigma_d$, in order to exclude fixed-position hot pixels and stellar interference that lack dynamic variation.

The thresholds $\mu_s+5\sigma_s$ and $\mu_d+3\sigma_d$ are empirically selected based on the characteristics of actual SLR videos. The spatial variance uses a stricter multiplier (5) to efficiently suppress single-frame random bright spots, background fluctuations, and localized noise. Conversely, the temporal variance utilizes a more relaxed multiplier (3) to avoid arbitrarily filtering out the genuine time-varying responses of weak targets. In practical applications, higher thresholds should be applied when background fluctuations are strong (to lower false alarm risk), while lower thresholds can be adopted for exceptionally faint targets where maximizing the detection rate is the primary objective. Only pixels that simultaneously satisfy the above two physical constraints are determined as valid target candidate regions. The fusion result is shown in Fig. 5.

2.5. Adaptive Mapping Based on the Noise Floor

Although the squaring operation in the algorithmic pipeline effectively highlights the target, it also introduces a signal masking effect. The intensity of strong laser signals often far exceeds that of weak satellites, causing traditional global normalization methods to lose weak target information—a phenomenon known as "darkness under the lamp". To address this, this paper proposes an adaptive mapping mechanism based on the background noise floor.



Fig. 5. Results of the dual-gate fusion mechanism, where only the clear target points remain in the entire frame.

The final output image, I_{out}, is constructed using a mapping logic that depends on the background standard deviation rather than the global maximum intensity:

$$I_{\text{out}}(x,y)=\min\left[\frac{S_{\text{fusion}}(x,y)}{k\cdot\sigma_{\text{floor}}},1\right], \tag{7}$$

where $S_{\text{fusion}}(x,y)$ denotes the intensity of the preserved target signal after the dual-gate screening, as defined in Equation (6); σ_{floor} denotes the background noise floor, represented by the standard deviation under the current observational conditions; and k denotes the dynamic saturation coefficient. In this study, k is set to 20, meaning that any target with a SNR greater than 20 is regarded as a high-confidence target.

The physical significance of this mechanism lies in shifting the focus from the absolute brightness of the most prominent features (such as overexposed laser signals) to whether a weak target stands out sufficiently against the background noise. Based on the analysis of real SLR videos, a threshold of $k=20$ is found to be optimal. Whenever the target intensity exceeds 20 times the background noise floor, the algorithm forcibly maps it to a high-brightness visible state (1.0). This approach effectively eliminates the "darkness under the lamp" effect caused by the coexistence of strong and weak signals.

3. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

To systematically evaluate the detection performance of the proposed algorithm in practical SLR engineering applications, this section designs a series of quantitative and qualitative experiments. The experiments are intended to verify the target identification capability and noise robustness of the algorithm under conditions of extremely low SNR and hot pixel interference.

3.1. Dataset Description

Data collection for this study was conducted within the operational environment of the TROS1000 system[16]. Developed by the Institute of Seismology, China Earthquake Administration, TROS1000 is currently the world's largest-aperture mobile SLR system. It is equipped with a 1m aperture optical telescope, features a maximum

ranging capability of 36000 km, and is deployed at the Nanshan Observatory of the Xinjiang Astronomical Observatory, Chinese Academy of Sciences.

For the performance evaluation, a total of 92 video sequences were selected. This dataset includes videos containing 12 different types of satellites (ranging from Low Earth Orbit to Geostationary Orbit), as well as "target-free" control group videos. It is important to clarify that these target-free videos are not simple dark-field calibration images; rather, they are actual observation segments that lack satellite targets (e.g., periods before the target enters the FOV, after it leaves, or pure sky background sequences). They retain the authentic background noise, sensor bad pixels, and stripe interference of real observation conditions, thereby serving as rigorous control samples to assess the algorithm's false alarm suppression capability.

3.2. Evaluation Metrics

To align with the practical requirements of automated SLR observations, this study adopts a binary decision standard centered on the "single observation mission".

In the experiments, each independent observation video is treated as an individual detection task. Given the spatio-temporal continuity of real satellite targets, the criterion for a valid detection is established. Specifically, a detection mission is judged as a "Success (Yes)" if the algorithm-extracted trajectory can continuously lock onto the target for more than N frames (set to $N = 30$ in this study); otherwise, it is judged as a "Failure (No)". Based on this criterion, we adopt two core engineering metrics, namely the Detection Success Rate (DSR) and the False Alarm Rate (FAR). The DSR measures the system's ability to capture the intended target and is defined as the ratio of videos in which the algorithm successfully identifies a valid target to the total number of videos containing real satellites. By contrast, the FAR measures the system's robustness against false positives. It is defined as the proportion of cases in which the algorithm incorrectly outputs a false target, such as locking onto hot pixels, in target-free background videos, during intervals when the target is absent, or when multiple targets are erroneously labeled, although only one target is tracked in a standard observation.

3.3. Computational Efficiency and Real-time Performance

To evaluate the real-time applicability of the proposed algorithm at SLR stations, we benchmarked its computational efficiency. The hardware configuration and quantitative metrics are summarized in Table 1.

Extensive evaluations over 6485 continuous frames demonstrate that the proposed algorithm achieves a steady-state throughput of 64.85 fps, robustly satisfying the strict 30 fps real-time processing requirement of SLR stations. While the temporal sliding window mechanism inherently introduces a one-time initial buffering delay of 3.33 s, the system maintains continuous and stable real-time performance once the buffer is established.

Table 1. Computational efficiency and real-time performance metrics

Metric	Specification/Result
Hardware platform	AMD CPU RTX 4060 GPU
Input resolution	640×480
Temporal Window Length (T)	100 frames
Average processing time	15.42 ms/frame
Steady-state throughput	64.85 fps
Initial buffering delay	3.33 s (100 frames)

3.4. Test Results

To comprehensively evaluate the performance of the proposed detection algorithm based on spatio-temporal statistical feature fusion, the 92 real SLR observation sequences were processed. These sequences encompass 11 types of satellite targets with diverse orbital altitudes and reflective characteristics—including AJISAI, STELLA, and BeiDou-3—as well as the "target-free" pure observation background data.

Table 2 presents the detailed testing results across the comprehensive dataset. Out of the 92 video sequences, the algorithm achieved an overall DSR of 95.6% (88/92). Notably, only three false alarms were generated throughout the entire testing process, corresponding to a FAR of 3.3%. These experimental results demonstrate that the algorithm can effectively overcome the influence of strong interference and low contrast, achieving high-precision localization and extremely low false-alarm detection for weak SLR targets.

To further validate the performance of the proposed algorithm under diverse observational conditions, four representative satellite targets with different orbital characteristics and signal qualities were selected for detailed analysis.

(1) **AJISAI (LEO)**

AJISAI orbits at an altitude of approximately 1500 km. In the observation sequences, the target is characterized by high intrinsic brightness accompanied by significant scintillation effects. As shown in Fig. 6, the algorithm successfully captures the target's trajectory despite its rapid intensity fluctuations.

(2) **STELLA (LEO)**

Similar to AJISAI, STELLA orbits at approximately 800 km. However, the FOV in this sequence contains active laser emissions, resulting in localized overexposed regions. As shown in the tracking results, although the laser beam creates a masking effect, it typically manifests as a diffuse structure that fails the spatial compactness constraint. Consequently, the proposed algorithm successfully isolates the localized target response without requiring direct laser removal during preprocessing (Fig. 7).

(3) **LAGEOS (MEO)**

LAGEOS orbits at an altitude of approximately 5900 km. The observation data for this satellite is characterized by an extremely low SNR, substantial hot pixel inter-

Table 2. Detailed detection results of the proposed algorithm, LEO:Low Earth Orbit; MEO: Medium Earth Orbit

Target type	NORAD ID	Orbit/Altitude	Total sequences	Success (DSR)	Failure	Remarks
AJISAI	16908	LEO (~1500 km)	11	11 (100%)	0	–
STELLA	22824	LEO (~800 km)	3	3 (100%)	0	–
BeiDou-3 (M14)	43623	MEO (~21500 km)	2	2 (100%)	0	–
BE-C	01328	LEO (~1000 km)	9	9 (100%)	0	–
CRYOSAT-2	36508	LEO (~700 km)	4	4 (100%)	0	–
GLONASS-142	45358	MEO (~19100 km)	1	1 (100%)	0	–
HY-2C	46469	LEO (~970 km)	7	6 (85.7%)	1	False alarm
LARETS	27944	LEO (~690 km)	3	3 (100%)	0	–
SENTINEL-3	41335/43437	LEO (~800 km)	4	4 (100%)	0	–
STARLETTE	07646	LEO (~800 km)	9	9 (100%)	0	–
LAGEOS	08820/22195	MEO (~5900 km)	4	3 (75%)	1	Missed detection
No target	–	–	35	33 (94.3%)	2	False alarm
Total	–	–	**92**	**88 (95.6%)**	**4 (4.3%)**	–

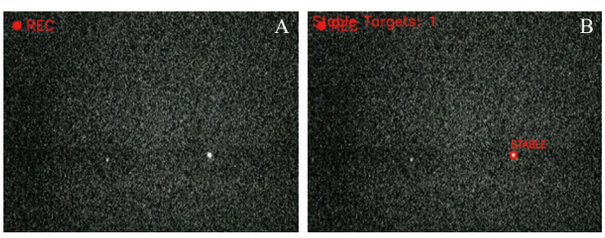


Fig. 6. Detection performance for AJISAI. (A) Original image. (B) Detection result with red marker indicating the target.

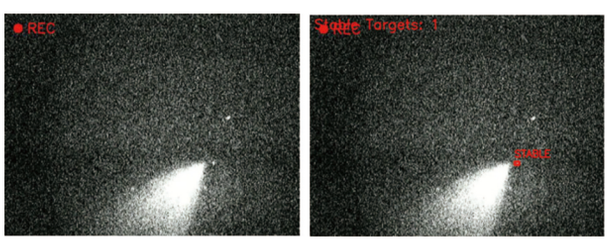


Fig. 7. Detection performance for STELLA under laser interference.

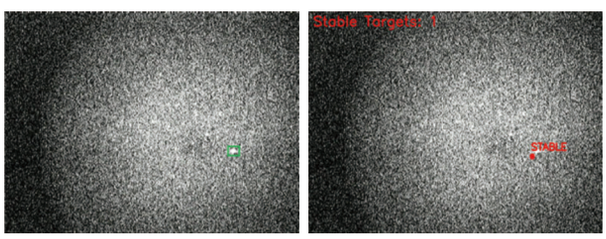


Fig. 8. Detection performance for LAGEOS showcasing clutter suppression.

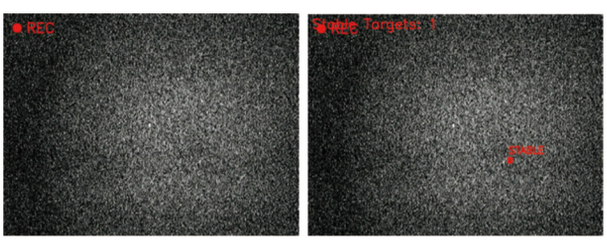


Fig. 9. Detection performance for the weak GLONASS-142 satellite signal.

ference, and the presence of other moving non-target objects within the FOV. In Fig. 8 (left), the bright spot within the green box represents a moving interference object (clutter) rather than the intended target. The dual-channel fusion mechanism successfully discriminates and excludes this interference, achieving precise target identification as shown in the right panel.

(4) **GLONASS-142 (MEO)**

GLONASS-142 orbits at an altitude of approximately 19100 km. Due to its high altitude, the target signal is exceptionally faint and is virtually indistinguishable from the background noise in a single frame. The algorithm leverages temporal-spatial accumulation to enhance the signal, enabling successful detection where traditional methods fail, as shown in Fig. 9.

4. CONCLUSION

To address the persistent challenges of low SNR and the difficulty of identifying targets in SLR video sequences, this paper presents a high-precision detection algorithm specifically tailored for SLR. Comprehensive evaluations conducted on a dataset of 92 real-world observation sequences—encompassing 11 satellite types across various orbital altitudes, such as AJISAI and LAGEOS—demonstrate that the proposed algorithm achieves an overall DSR of 95.6% (reaching 100% in specific scenarios) with a FAR of only 3.3%. The experimental results and practical deployment confirm that the fusion of spatio-temporal statistical features effectively suppresses complex interferences, including sensor-intrinsic noise, dynamic clutter, and strong signal masking. Notably, the algorithm was successfully operationalized during the 2025 Xinjiang SLR observation.

AI DISCLOSURE STATEMENT

AI-assisted technology is not used in the preparation of this work.

AUTHOR CONTRIBUTIONS

Jinlong Hu conceived the study, developed the methodology, conducted the experiments, analyzed the data, and drafted the manuscript. Yi Zhang contributed to the experimental design, data analysis, and manuscript revision. Tong Zou supervised the research, provided guidance on the methodology and result interpretation, and revised the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

[1] Lin, H. S., Wu, Z. B., Zheng, M., et al. 2023. Research and application of picosecond accuracy time delay calibration for satellite laser ranging system. *Infrared and Laser Engineering*, **52**(10): 106−117. (in Chinese)

[2] Thaller, D., Dach, R., Seitz, M., et al. 2015. Satellite laser ranging to GPS and GLONASS. *Journal of Geodesy*, **89**: 535−535.

[3] Wang, X., Xiao, W., Li, W. C., et al. 2024. Progress of Xi'an satellite laser ranging system of NTSC. *Journal of Time and Frequency*, **47**(4): 268−276. (in Chinese)

[4] Zeng, M., Li, J. X., Peng, Z. 2006. The design of top-hat morphological filter and application to infrared target detection. *Infrared Physics & Technology*, **48**(1): 67−76.

[5] Chen, C. L. P., Li, H., Wei, Y., et al. 2014. A local contrast method for small infrared target detection. *IEEE Transactions on Geoscience and Remote Sensing*, **52**(1): 574−581.

[6] Wei, Y. T., You, X. G., Li, H. 2016. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognition*, **58**: 216−226.

[7] Gao, C. Q., Meng, D. Y., Yang, Y., et al. 2013. Infrared patch-image model for small target detection in a single image. *IEEE Transactions on Image Processing*, **22**(12): 4996−5009.

[8] Stauffer, C., Grimson, W. 1999. Adaptive background mixture models for real-time tracking. In Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR).

[9] Helferty, J., Smith, J., Ryan, E., et al. 2023. Performance analysis of satellite tracking algorithms in low snr environments. In Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS).

[10] Delaite, T., Couetdic, J., Glémet, É., et al. 2023. Performance of an Optical COTS Station for the Wide-Field Detection of Resident Space Objects. In Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS).

[11] Salmond, D., Birch, H. 2001. A particle filter for track-before-detect. In Proceedings of the 2001 American Control Conference (ACC).

[12] Reed, I., Gagliardi, R., Stotts, L. 1988. Optical moving target detection with 3-D matched filtering. *IEEE Transactions on Aerospace and Electronic Systems*, **24**(4): 327−336.

[13] Min, F. Q., Zou, T., Guo, T. Y., et al. 2019. A target recognition method of SLR system based on image processing. *Beijing Surveying and Mapping*, **33**(4): 433−437. (in Chinese)

[14] Wang, Y. J., Guo, T. Y., Zou, T., et al. 2013. Study of Recognition Techniques of SLR Target Based on Image Accumulation. *Journal of Geodesy and Geodynamics*, **33** (Suppl 2): 71−73. (in Chinese)

[15] Cheng, H. Y., King, Y. C. 2003. A CMOS image sensor with dark-current cancellation and dynamic sensitivity operations. *IEEE Transactions on Electron Devices*, **50**(1): 91−95.

[16] Guo, T. Y., Wang, P. Y., Li, X., et al. 2015. Progress of the satellite laser ranging system TROS1000. *Geodesy and Geodynamics*, **6**(1): 67−72.