Advancing Space Situational Awareness: Detecting Resident Space Objects with Neuromorphic Vision Sensor

Yusra Alkendi[1†], *Alexey Simonov*[1], *Konstantin Kravtsov*[2], *Sana Amairi-Pyka*[2], *James A. Grieve*[2], *and Anton B. Ivanov*[1]

[1]*Propulsion and Space Research Center (PSRC),* [2]*Quantum Research Center (QRC), Technology Innovation Institute (TII), Masdar City, Abu Dhabi, United Arab Emirates*

Emails: yusra.alkendi@tii.ae, alexey.simonov@tii.ae, konstantin.kravtsov@tii.ae, sana.pyka@tii.ae, james.grieve@tii.ae, anton.ivanov@tii.ae

[†]*Corresponding author*

Abstract

Space Situational Awareness (SSA) involves detection, tracking, and forecasting the movements of objects in Earth's orbit. It is crucial for protecting space assets and preventing collisions, yet challenging due to the vast number of objects, their high-velocity interactions, and difficult lighting conditions. In this project, we propose using a Neuromorphic Vision Sensor (NVS) to observe and detect Resident Space Objects (RSOs) to advance SSA field. The NVS is an innovative, bio-inspired sensor that asynchronously triggers logarithmic light intensity changes at the pixel level. The sensor will be mounted on a high-end 0.8 m diameter Ritchey–Chrétien telescope, part of the Abu Dhabi Quantum Optical Ground Station (ADQOGS) equipped with less than 8" pointing accuracy. This NVS event camera is activity-dependent, resulting in significantly less data for sparse scenes compared to traditional cameras. Furthermore, the NVS can capture high-speed motion and fine details of RSOs, overcoming the limitations of conventional cameras that suffer from motion blur and low resolution. By exploiting the unique characteristics of the NVS, we aim to develop a novel and robust approach for RSO detection that can handle various scenarios and lighting conditions. Coupled with reduced power consumption, lower processing requirements, wider dynamic ranges, and faster communication speeds, these attributes make NVS exceptionally well-suited for space imaging and SSA applications. We are developing an innovative deep learning algorithm to process NVS event stream observations, enabling it to differentiate between event streams related to the sky background and those related to RSOs. The algorithm processes and reveals the spatial-temporal correlations between events in the NVS data and outputs the event labels. The algorithm is trained on publicly available NVS-based observations, and initial evaluations have shown promising results, demonstrating the potential of Artificial Intelligence (AI) for processing NVS raw data in SSA applications. Moving forward, we aim to contribute to the advancement of space imaging technology while ensuring the reliability and validity of our research findings. The deliverables include the experimental setup equipping NVS sensor to a local ground telescope and calibration, the advanced deep learning algorithms for RSO detection, tracking, and classification and the creation of a new NVS-based dataset recorded locally in the UAE. These efforts are poised to substantially elevate the UAE's capabilities in RSO monitoring, contributing to enhanced accuracy, efficiency, and operational insights, thereby strengthening space situational awareness and the overall management and research of space objects.

Keywords: Neuromorphic Vision Sensors, Space Situational Awareness, Artificial Intelligence, Deep Learning, Resident Space Objects

1. Introduction

Space Situational Awareness (SSA) is critical for ensuring the safety and sustainability of space operations, particularly in detecting, tracking, and mitigating the risks posed by space debris. As the number of objects in Earth's orbit continues to grow, accurate and efficient debris detection is crucial for avoiding collisions and protecting valuable space assets. Traditionally, image-based sensors such as Complementary Metal-Oxide-Semiconductor (CMOS) cameras have been vital for spacecraft operations, astronomy, and remote sensing. However, these systems face limitations in handling the high velocities and challenging lighting conditions prevalent in space environments.

Neuromorphic Vision Sensors (NVSs) present a transformative advancement in SSA, offering an alternative to conventional sensors. NVSs utilize asynchronous event-triggering based on logarithmic light intensity changes at the pixel level, providing significant advantages over frame-based technologies like CMOS cameras. Each pixel in an NVS functions autonomously, generating data solely when a contrast change is detected, allowing the sensor to react swiftly and dynamically to changes in its environment

The unique in-pixel circuitry of NVSs allows for high-speed, low-power, and high-dynamic-range imaging. Unlike

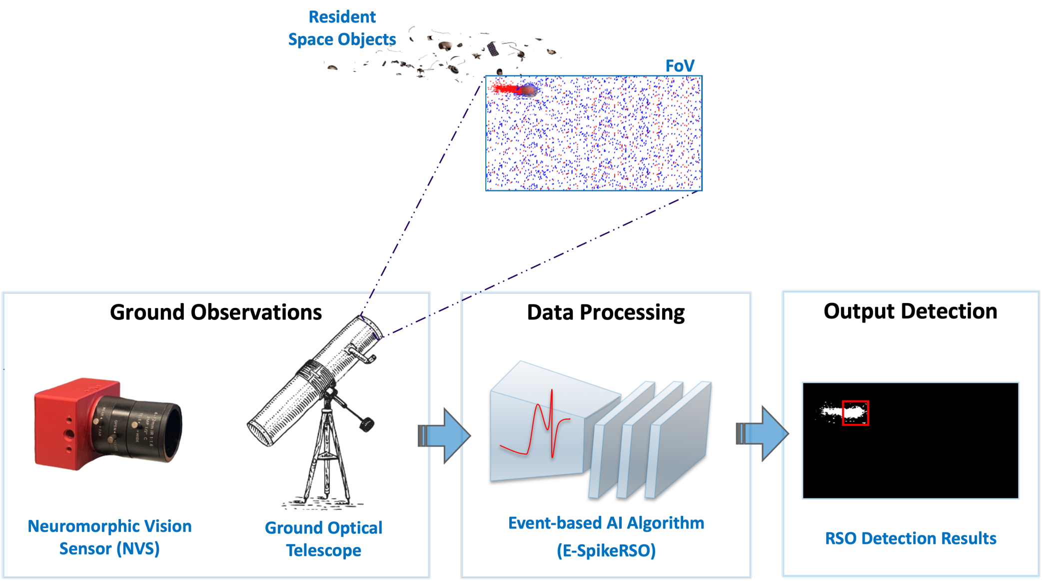

Figure 1: Proposed Framework - Neuromorphic Vision Sensor (NVS) for Space Situational Awareness (SSA).

traditional sensors that rely on fixed-time integrations, NVSs generate a continuous stream of change events with microsecond precision. Each event contains information about the pixel's location, the direction of change, and the precise time it occurred, represented as a tuple of (x, y, t, p), where (x, y) denotes the pixel's location, t is the event time, and p is a binary value indicating an increase or decrease in luminance. As NVSs produce data only when a change occurs, they generate significantly less data for sparse scenes compared to conventional sensors, making them highly efficient for space imaging applications [1, 2].

The unique properties of NVSs have driven researchers to explore their potential in space applications. NVS technology has been investigated for tasks such as spacecraft navigation and control, particularly in time-to-contact predictions for space landings [3, 4], as well as for star and satellite tracking [5, 6]. Moreover, NVSs have been employed for space astrometry and the detection of resident space objects (RSOs), contributing to significant advancements in SSA [7, 8, 9, 10].

In this work, we propose leveraging NVS technology coupled with AI-based algorithms to improve RSO detection and tracking at the Abu Dhabi Quantum Optical Ground Station (ADQOGS), as depicted in Fig. 1. By leveraging the high-speed, low-power, and high-dynamic-range capabilities of NVS, alongside advanced deep learning models, we aim to develop a scalable module capable of addressing the challenges of space debris detection. We aim to contribute advancing space imaging technologies, providing a more reliable and efficient solution for SSA.

The remainder of this paper is structured as follows: Section 2 presents the problem statement and contributions of the work. Section 3 details the proposed AI-based algorithm for space object detection. In Section 4, we provide a thorough experimental evaluation, including both quantitative and qualitative analyses. Conclusions drawn from this research are discussed in Section 5.

2. Problem Statement and Contribution

This research project focuses on developing and implementing a debris detection and tracking module using a resource-constrained onboard computer through a ground-based telescope (ADQOGS). By leveraging the innovative *Neuromorphic Vision* technology, specifically designed for *Space Situational Awareness (SSA)*, the module will address the growing challenges of space debris management. The importance of this work lies in the protection of space assets, the mitigation of space traffic, and fostering international collaboration. Additionally, the project will involve the collection of a local dataset, further enhancing research capabilities. Ultimately, this effort aims to advance the UAE's technological expertise in space exploration, contributing to both national and global SSA initiatives. In this work, we design and develop an AI-based algorithm as the initial step of the project, where we investigate the potential of AI coupled with the NVS camera for space object detection. This will help explore potential processing algorithms for object detection, thereby advancing Space Situational Awareness (SSA), as will be discussed in Section 3.

3. Proposed Framework Neuromorphic Vision Sensor (NVS) for Space Situational Awareness (SSA)

In this section, the proposed framework is divided into three stages: (1) Data Processing - E-SpikeRSO Algorithm for Space Object Detection, (2) Data Processing - E-SpikeRSO Algorithm for Space Object Detection, and (3) Output Detection of the Object of Interest. Each stage is presented in detail in the following subsections.

3.1. Data Acquisition - Ground Observations

ADQOGS is located in Adu Dhabi at the elevation of 70 m above sea level at 24°11′ N, 54°41′ E. Its main instrument utilized for data acquisition is a Ritchey-Chrétien telescope with

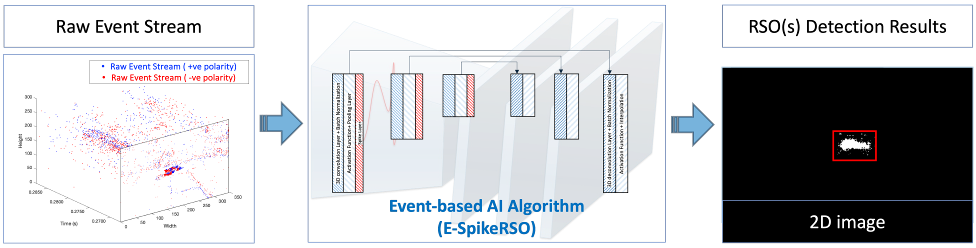

Figure 2: Framework of our E-SpikeRSO-based network for detecting RSOs from NVS event streams.

the mirror aperture of 80 cm at f/6.85. The telescope's location is a compromise that provides good connectivity to the city's optical data networks, while also being far enough from the city to enhance visibility conditions. The site infrastructure allows for fully-automated telescope operation, so SSA data acquisition may be performed at the idle time of the instrument, when it is not used for other tasks. The NVS camera will be placed at one of the two available Nasmyth ports by means of specially designed adapters mounted on the telescope body.

3.2. Data Processing - E-SpikeRSO Algorithm for Space Objects Detection

In this section, we will discuss the input data preparations, the architecture of the proposed *E-SpikeRSO* model, the approximate ground truth data preparation, and the training and testing methodology for *E-SpikeRSO*.

3.2.1. Input Data Preparations

The data captured by a NVS event camera records changes in the logarithmic intensity of a scene as asynchronous events. These events occur across a spatial resolution of $H \times W$ pixels, where H and W represent the height and width dimensions of the camera's frame, respectively. A stream of N events, denoted by $\{e_i\}_N$, is described as a sequence of 4-tuples:

$$\{e_i\}_N = \{x_i, y_i, t_i, p_i\}_N, \tag{1}$$

where (x_i, y_i) are the pixel coordinates of the i-th event, t_i is the timestamp, and p_i indicates the event's polarity. The polarity p_i takes a value of $+1$ if the pixel brightness increases and -1 if it decreases.

To handle these events, they are aggregated into a 3D-event tensor $\{x_i, y_i, t_i\}_N$ over a predefined temporal window excluding the polarity p_i. In low-light conditions or when capturing scenes with dark backgrounds, such as the night sky, the high temporal resolution of event cameras can produce a substantial amount of simultaneous noise events. This leads to significant data redundancy and increases computational complexity, making it challenging to accurately interpret the spatiotemporal relationships between events.

To mitigate these issues, the 3D-event tensor is constructed every 5ms, reducing the amount of data processed. These tensors are then passed through the encoding-decoding layers of the proposed E-SpikeRSO algorithm, which outputs a 2D image that segments the object of interest, as described in Section 3.2.2.

3.2.2. Architecture of the Proposed E-SpikeRSO Model

Figure 2 shows the overall architecture of E-SpikeRSO, where the compiled 3-event tensor is used as input to the network. The active pixel locations and the associated time features from the event stream are processed by encoding and decoding layers to detect the object of interest within the frame. The network outputs a 2D image in which each pixel is classified as either part of the sky background or an RSO object.

Each encoder unit consists of a convolutional layer followed by a Spike layer, allowing the network to process varying sizes of the 3D-event tensor as it passes through. On the other hand, the decoder units consist solely of deconvolutional layers to produce the 2D output image. The number of encoder and decoder units can be adjusted based on the specific application and data characteristics. In this case, the network architecture includes three encoder-decoder units, as shown in Fig. 2. The final output is passed through a softmax layer, assigning each pixel the predicted probability of being classified as either an RSO object or part of the sky background.

3.2.3. Approximate Ground Truth Data Preparation

The proposed E-spikeRSO is developed, trained, and evaluated using publicly available datasets from [9], known as the first event-based space imaging dataset (EBSSA). This dataset, accessible via the EBSSA Drive link, contains recorded event streams and corresponding masked RSO frames captured from multiple remote locations. Each entry in the dataset includes event streams paired with masked image frames of detected RSO objects.

For each 5 ms temporal window, the input to the network is a 3D-event tensor generated from all event data within that period. The corresponding output is a binary image, where pixels with a value of 1 represent the masked object in the scene, and pixels with a value of 0 represent the sky-background. A sample of a labelled event stream is shown as the network output in Fig. 2. After preparing the dataset samples, the network was trained on ten different sequences and tested on two unseen sequences that were excluded from the training process. The dataset features various event streams recorded under different conditions, including both stationary and moving telescopes. The algorithm is trained to detect not only RSOs but also streaks and stars in scenes where the telescope is in motion.

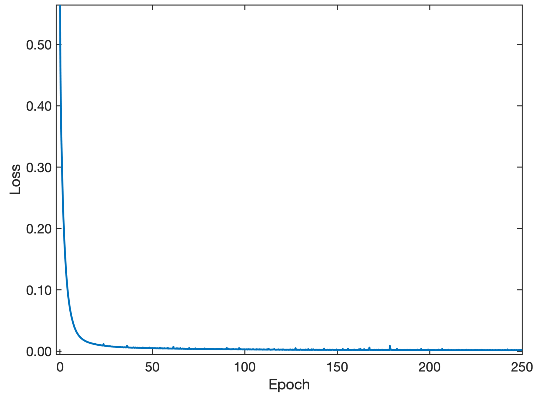

Figure 3: Loss curves generated during the training of the E-SpikeRSO Algorithm.

3.2.4. Training and Testing Methodology for E-spikeRSO

Our method involves training an E-SpikeRSO neural classifier on 3D-event tensors, which represent NVS event streams, to perform object detection in a supervised manner. The network outputs a 2D image that segments the approximate ground truth pixels corresponding to the object of interest. The sparse and limited features of the event data, represented as $< x_i, y_i, t_i, p_i >$, pose challenges for scene interpretation and precise object detection. Despite this, the high temporal resolution of event data, along with noise introduced in low-light conditions, facilitates the rapid accumulation of large volumes of event information.

The performance of AI-based models heavily depends on the quality and size of the datasets used during training and testing. Providing neural networks with a wide and diverse range of training data is crucial for two key reasons: (I) fine-tuning the model's parameters for accurate predictions, and (II) improving its ability to generalize to new, unseen data. However, acquiring and accurately annotating large datasets can be challenging and resource-intensive for certain applications. The network is implemented in PyTorch for both the training and testing phases, utilizing the Adam optimizer with a learning rate of 0.001 to minimize cross-entropy loss. Figure 3 shows the loss curves produced by the proposed training scheme, demonstrating the model's convergence.

3.3. Output Detection

The E-SpikeRSO algorithm outputs a 2D segmented image that classifies each pixel as either part of the object of interest (RSO) or the sky background. By processing 3D-event tensors from NVS streams, the network accurately predicts the approximate location of RSOs, generating a binary mask where pixels with a value of 1 represent the RSO and 0 represent the background. Despite challenges posed by noise and sparse event data, the network's architecture effectively captures spatiotemporal correlations to detect RSOs in various conditions. In the next section, we will evaluate the model both quantitatively and qualitatively. In the next section, quantitative and qualitative evaluations of the model are conducted, highlighting the potential of AI-based algorithms for reliable RSO detection applications.

4. Experimental Evaluations

In this section, the experimental evaluations of the proposed *E-SpikeRSO* model is presented. We begin by introducing the evaluation metrics used to assess the model's performance, followed by a detailed discussion of both quantitative and qualitative analyses of the results.

4.1. Evaluation Metrics

To quantitatively evaluate the performance of the proposed model, we focus on two key metrics: *Precision* and *Accuracy*, which are commonly used in classification tasks.

- *Precision*: Precision assesses the proportion of correctly detected objects (true positives) out of all detected objects (true positives and false positives), highlighting the model's ability to minimize false positives. It is defined as:

$$\textit{Precision} = \frac{TP}{TP + FP} \tag{2}$$

- *Accuracy*: Accuracy represents the ratio of correctly classified active pixels (both object-related and background-related) to the total number of active pixels within the frame, providing an overall measure of the model's classification performance. It is expressed as:

$$\textit{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

where, TP, FP, TN, and FN stand for true positives, false positives, true negatives, and false negatives, respectively. True positives refer to correctly detected objects, while true negatives indicate correctly identified background events.

4.2. Quantitative and Qualitative Analyses

Table 1 summarizes the performance of the proposed E-SpikeRSO algorithm, assessed on both training and testing sequences using the *Precision* and *Accuracy* metrics. It is worth noting that *E-SpikeRSO* is a binary image segmentation model trained on highly imbalanced data, where the number of active pixels corresponding to the object of interest is considerably smaller than those representing the sky background. To tackle this challenge and the limited input features from the event stream, we integrated a Spike layer with a convolutional layer in the encoding units. This combination effectively captures the sparse and asynchronous nature of event-based while retaining important spatiotemporal information. The Spike layer emulates the way biological neurons process information, enhancing the model's ability to handle sparse activity in the input, while the convolutional layer extracts essential features to improve pixel-wise classification.

Table 1: Evaluation Metrics for Training and Testing Sets

	TP	TN	FP	FN	Precision	Accuracy
Training sets	57311	22011314	209	6366	99.6366%	99.9702%
Testing sets	903	1162069	164	3264	84.6298%	99.7061%

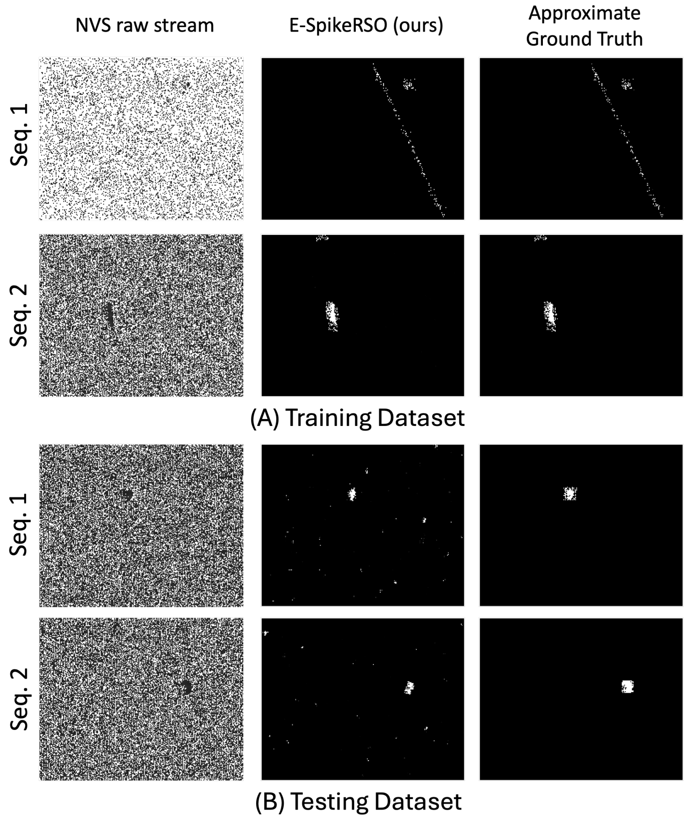

Figure 4: RSO Detection results using our E-SpikeRSO algorithm on NVS event streams during training and testing sets.

Table 1 presents that the model performs similarly on both unseen test data and training data across various evaluation metrics. Due to the dataset's imbalance, the model shows notably high *Accuracy*. However, a thorough assessment requires not only considering *Precision* and *Accuracy*, but also incorporating qualitative analysis for a more balanced view. Figure 4 presents sample visualizations comparing the model's predictions on training and test data with the approximated ground truth. Both the qualitative and quantitative analyses indicate that the *E-SpikeRSO* algorithm performed well as an initial model. This demonstrates that using a combination of convolutional and Spike layers offers a promising models for developing NVS-based detection modules for Resident Space Objects (RSOs).

5. Conclusions

In this work, we demonstrate the potential of Neuromorphic Vision Sensors (NVS) coupled with an AI-based algorithm as an advanced technology for enhancing Space Situational Awareness (SSA). Our AI algorithm, the E-SpikeRSO model, successfully distinguishes RSOs from background noise in NVS event streams, showing promising preliminary results. Moving forward, the development of a locally recorded NVS-based dataset in the UAE, using the Abu Dhabi Quantum Optical Ground Station (ADQOGS), will further enhance our model's performance and adaptability. The successful implementation of this technology is poised to substantially elevate the UAE's SSA capabilities, ensuring greater accuracy, efficiency, and operational insights in space object management. Ultimately, our research contributes to the advancement of space imaging technologies, reinforcing the foundation for safer and more informed space operations

References

[1] X. Zheng, Y. Liu, Y. Lu, T. Hua, T. Pan, W. Zhang, D. Tao, L. Wang, Deep learning for event-based vision: A comprehensive survey and benchmarks, arXiv preprint arXiv:2302.08890 (2023).

[2] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis, et al., Event-based vision: A survey, IEEE transactions on pattern analysis and machine intelligence 44 (1) (2020) 154–180.

[3] S. McLeod, G. Meoni, D. Izzo, A. Mergy, D. Liu, Y. Latif, I. Reid, T.-J. Chin, Globally optimal event-based divergence estimation for ventral landing, in: European Conference on Computer Vision, Springer, 2022, pp. 3–20.

[4] O. Sikorski, D. Izzo, G. Meoni, Event-based spacecraft landing using time-to-contact, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 1941–1950.

[5] T.-J. Chin, S. Bagchi, A. Eriksson, A. Van Schaik, Star tracking using an event camera, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019, pp. 0–0.

[6] S. Bagchi, T.-J. Chin, Event-based star tracking via multiresolution progressive hough transforms, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2020, pp. 2143–2152.

[7] G. Cohen, S. Afshar, A. Van Schaik, Approaches for astrometry using event-based sensors, in: Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference, 2018, p. 25.

[8] G. Cohen, S. Afshar, B. Morreale, T. Bessell, A. Wabnitz, M. Rutten, A. van Schaik, Event-based sensing for space situational awareness, The Journal of the Astronautical Sciences 66 (2019) 125–141.

[9] S. Afshar, A. P. Nicholson, A. Van Schaik, G. Cohen, Event-based object detection and tracking for space situational awareness, IEEE Sensors Journal 20 (24) (2020) 15117–15132.

[10] N. Ralph, D. Joubert, A. Jolley, S. Afshar, N. Tothill, A. Van Schaik, G. Cohen, Real-time event-based unsupervised feature consolidation and tracking for space situational awareness, Frontiers in neuroscience 16 (2022) 821157.