General purpose, software configurable, intelligent LiDAR sensor for space-based non-cooperative resident space object relative navigation and tracking applications

Joy Shohdy, Dr. Robert Karl, Jr., Brad Short, Dr. Ricardo Delgadillo, Bruce Anderson, Alex Sandoval-Mathamba, Michael Dahlin
Advanced Scientific Concepts LLC

ABSTRACT

A key challenge in real-time space situational applications is the ability of an in-space sensor to detect, classify, and track non-cooperative resident space objects (RSOs). Crucial real time parameters required for autonomous threat assessment and response include relative position, orientation, and 3D velocity between the in-space sensor and the RSO. Conventional situational awareness sensor suites composed of visible / IR imaging and scanning LIDAR sensors lack the precision necessary for timely and accurate RSO assessment data and typically have high Size, Weight, and Power (SWaP) penalties.

ASC's Global Shutter Flash LiDAR (GSFL) addresses this challenge. A compact solid-state LiDAR camera, the GSFL captures range and intensity in organized point cloud data format. The camera has been involved with both in-orbit and deep space operations, earning it a Technology Readiness Level 9 (TRL-9). The space-qualified GSFL can be adapted for various distances, making it suitable for real-time space situational awareness tasks like Space Domain Awareness, Debris Removal, Satellite Servicing, Characterization, and Tracking of both Non-Cooperative and Cooperative Threats.

These various applications can be accommodated through just two main camera configurations, each designed to fit within manageable SWaP budgets. Each configuration is customized with a unique field of view and laser divergence, optimized to fulfill specific objectives. The first is a short-range option (150 m to 5 km) suitable for Servicing and Proximity Operations. The second is a long-range option (1.3 km to 60 km) tailored for Debris Removal and Characterization. These configurations can also cater to intermediate and far-range imaging, primarily optimized for Tracking and Classification. The paper delves into the optimization of each configuration for its intended use and presents expected performance outcomes.

1. INTRODUCTION

Global Shutter Flash LiDAR (GSFL) employs a time-of-flight detector array and an optical diffuser to image an entire field of view in a single pulse, eliminating point cloud blurring or distortion. Further, since the data is collected as an array, there is no need for any post-processing to create an organized point cloud; the raw data is already organized. These features are vital for real-time applications. The structured and correlated nature of organized point cloud data is ideal for artificial intelligent / machine learning (AI/ML) processing. The GSFL has a 128x128 focal plane array (FPA) which is suitable for most applications, but due to the fixed pixel grid, the GSFL data can be easily fused with an electro-optical (EO) sensor to create even higher resolution images and to support advanced AI/ML algorithms. Recent research [3] focuses on noise filtering and AI/ML algorithms for target detection, classification, and tracking. These AI/ML algorithms can be reconfigured on-orbit for optimal performance in specific RSO scenarios.

This paper explores Advanced Scientific Concept’s (ASC) Global Shutter Flash LiDAR's capability to address the challenge of real-time space situational awareness. By adapting the camera to various configurations and integrating AI/ML algorithms, it offers precise and timely data for tasks like RSO tracking, classification, and threat assessment. The research emphasizes optical optimization, performance expectations, and data fusion potential, showcasing flash LiDAR’s potential for space applications.

2. SYSTEM PERFORMANCE CALCULATIONS

The goal of the optical optimization is to design a system that is operational over a wide span of ranges to target without a need to reconfigure the hardware. The broader the coverage range of each configuration, the fewer hardware variations are needed to cater to a diverse set of space applications. The capabilities of these system configurations are calculated through an optical link budget model assuming a 2-meter square target with 30% Lambertian reflectivity. For this study, a 60mJ, 1064 nm laser is assumed. The f-number of the receiving lens was taken to be 2.8. The primary optical tradeoff is in the lens field of view (FOV) and laser divergence; with increased divergence and FOV, a wider area is measured, but at a shorter operational range, whereas at a narrower divergence and FOV, the operational range is increased with a reduced field of regard. Two main camera configurations are presented: a short-range option and a long-range option.

Each configuration has a fixed field of view and four laser divergence settings. This leads to four modes of operation for each configuration. An actuated diffuser wheel allows for adjusting laser divergence by changing the optical diffuser in the path of the laser, permitting a smooth transition between modes of operation.

2.1 CONFIGURATION 1: SHORT-RANGE IMAGING

The Configuration 1 specifications, designated as the short-range configuration, are summarized in Table 1. It presents the calculated minimum and maximum operational range and the number of illuminated pixels for each operational mode. For these calculations, the maximum operational range is defined as the distance at which the predicted probability of detection for a single pixel falls below 95% [2]. The minimum range is defined as the point at which the sensor pixel enters saturation; at this distance, the LiDAR still reports range data, just with lower precision. The specified pixel count represents the quantity of pixels that sample the region illuminated by the laser within this divergence angle. Pixels beyond this region do not capture any returning light and do not require reading by the LiDAR. This configuration, optimized for Servicing, Proximity Operations, Tracking, and Classification tasks, is suitable for operational distances spanning from 150 meters to 5 kilometers.

Table 1. Configuration 1: Short-range

Operational Mode	Lens FOV (deg)	Laser Divergence (deg)	Min Range (m)	Max Range (m)	Pixels
1.1	15	15	150	650	128x128
1.2	15	6	380	1270	51.2x51.2
1.3	15	1.5	1220	2530	12.8x12.8
1.4	15	.375	2450	5060	3.2x3.2

The expected performance of this configuration is also presented in Fig. 1. This figure illustrates the photons per pixel per pulse on the sensor as a function of range to target for each operational mode. There are range values for which multiple operational modes are capable of proper functionality. This allows for smooth transitions between different operational modes and eliminates the possibility of losing a target as the system transitions between laser divergence settings. This means the LiDAR has no blind spots as it shifts from one operational mode to the next.

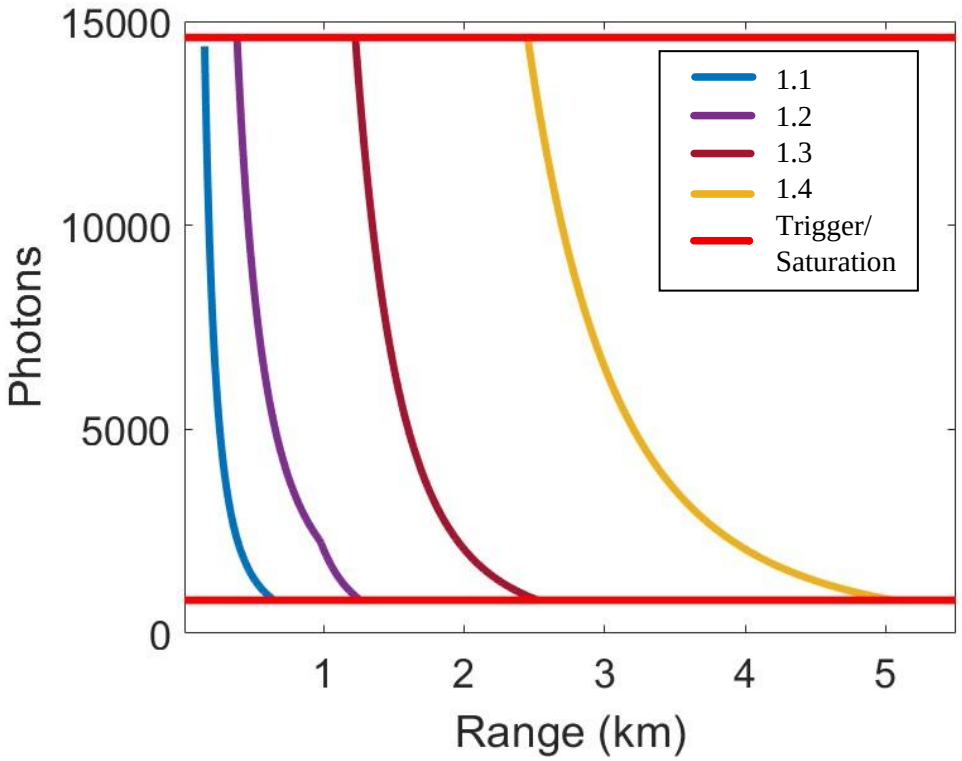


Fig. 1. Configuration 1: Range vs. Photons per Pixel

2.2 CONFIGURATION 2: LONG-RANGE IMAGING

Configuration 2 is optimized for long-range imaging. Its specifications and expected performance are summarized in Table 2. The definitions of the values reported are the same as for the short-range configuration. This configuration is suitable for Debris Removal and Characterization. It can also be tailored for intermediate to long-range Tracking and Classification applications. Small diffuser angles and a beam expander can permit imaging ranges from 1.3 km to 60 km, allowing for transition from intermediate to far-range imaging applications.

Table 2. Configuration 2: Long-range

Operational mode	Lens FOV (deg)	Laser divergence (deg)	Min Range (km)	Max Range (km)	Pixels
2.1	1.5	1.5	1.3	5.9	128x128
2.2	1.5	.4	5.1	14.7	34.1x34.1
2.3	1.5	.1	14.2	29.4	8.5x8.5
2.4	1.5	.024	29	60	2.0x2.0

Similar to Configuration 1, there is an overlap between the imaging range of each operational mode to allow for transition between laser divergences without causing blind spots. This is illustrated in Fig. 2 which shows the expected photons per pixel vs. imaging range for the sub-configurations of the long-range system.

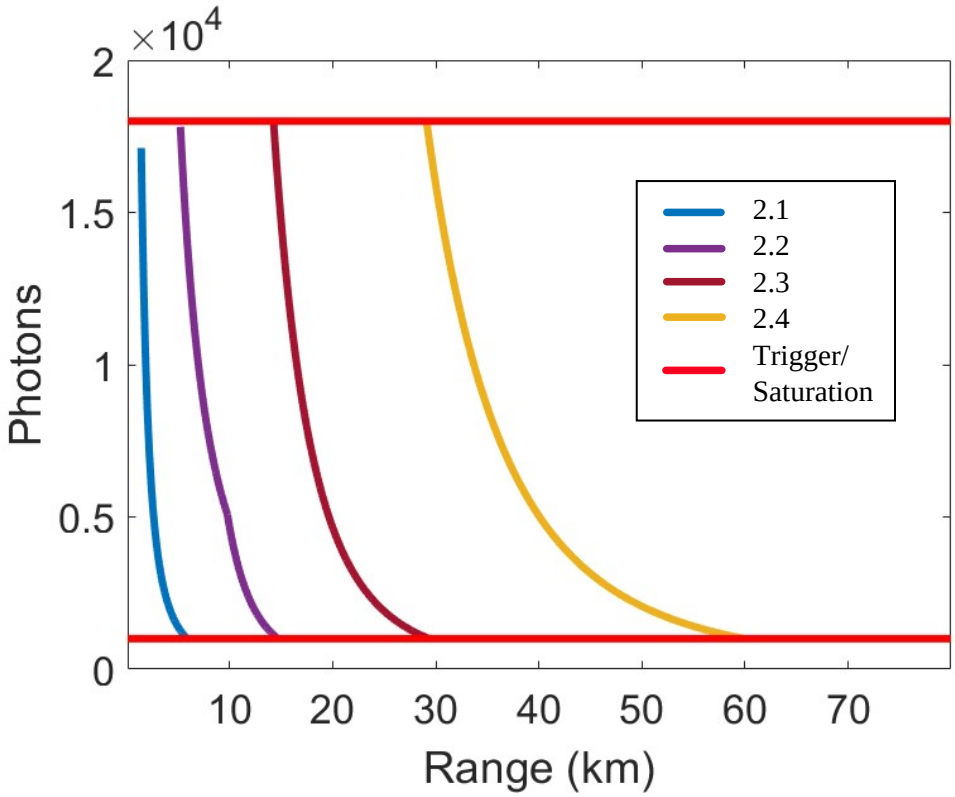


Fig. 2. Configuration 2: Range vs. Photons per Pixel

3. MECHANICAL DESIGN CONCEPT

Advanced Scientific Concepts' mechanical design concept for a configurable GSFL camera consists of two parts: an optical head and a processing module. The camera components are divided into two enclosures. The optical enclosure houses all optical components, laser, and sensors, while the processing module houses the main processing and electrical boards. The two enclosures are connected by radiation tolerant electrical cables of lengths that can be set by application. Configurability and mounting options are maximized for the system by separating the enclosures for the optical and processing components.


Fig. 3. (Left) Rendered Image of Configuration 1. (Right) Rendered Image of Configuration 2.

Fig. 3 illustrates the two configurations of the camera. The short-range configuration requires the 15° field of view lens while the long-range option requires the 1.5° lens. The entire camera system is designed to fit within a practicable SWaP budget. The optical enclosure has size of 9.75in x 10.25in x 6in. The processing box has a an even smaller size of 6.75in x 8.5in x 4.2in. Fig. 4 displays an exploded view of the short-range camera configuration and details of the components in each enclosure.

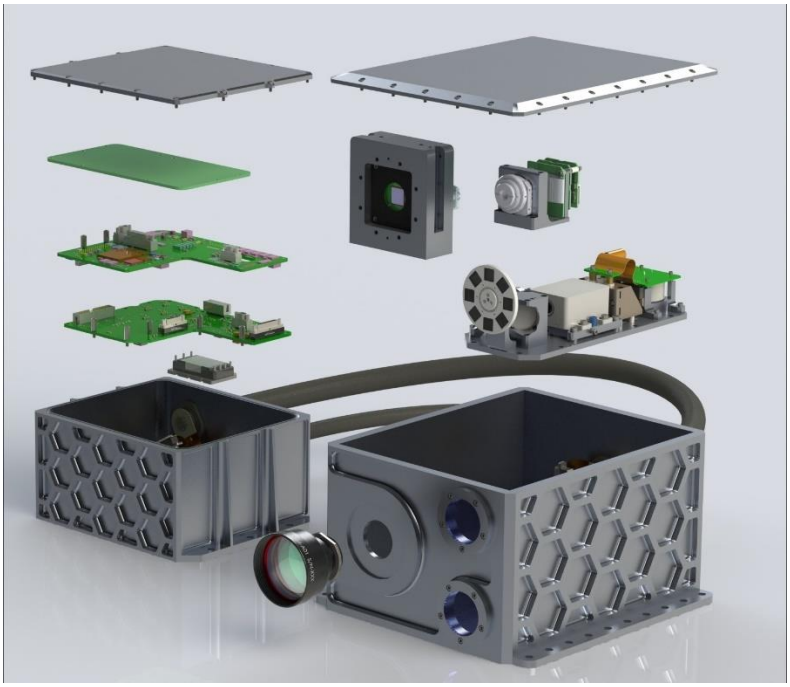


Fig. 4. Exploded view of Configuration 1: Short-range

4. SOFTWARE CONFIGURABILITY USING MACHINE LEARNING ALGORITHMS: SEMANTIC SEGMENTATION

4.1 RSO SEMANTIC SEGMENTATION

Hardware configurability can optimize the camera for certain range targets, but software configurability allows the camera to be even further optimized for specific space applications. ASC is developing ML/AI algorithms for optimal performance in RSO scenarios. ASC's current research has a strong emphasis on Semantic Segmentation for RSOs through the application of Deep Neural Networks (DNNs).

The goal for RSO DNN Semantic Segmentation is to be able to gain a better understanding of an RSO by predicting a meaningful class label for each data point collected by the GSFL. Semantic Segmentation is explored using simulated data of the ASC GSFL camera and using the Semantic Segmentation DNN known as SalsaNext [6]. SalsaNext is chosen as the Semantic Segmentation Neural Network as it performs well in the mean Intersection-Over-Union (mean IoU) metric and has a fast runtime. For numerical simulations, the data input to SalsaNext is of size 640kB, and the speed at which this DNN can process the data is around 20 milliseconds using the RTX 4090 GPU. With this speed, one can achieve a high frame rate real-time Semantic Segmentation.

4.2 SIMULATING DATA

To generate the necessary training data for SalsaNext, Computer Aided Design (CAD) models of Resident Space Objects (RSOs) are employed. NASA offers accessible CAD models through an online source, as referenced in [5]. The training and evaluation of SalsaNext will be conducted using a model of the TDRS-B satellite.

4.2.1 PREPARING THE LABELED CAD MODEL

The model is made up of a union of triangles, each triangle is described by an ordered triple of points

$$P1 = \{(x_i^1, y_i^1, z_i^1) | \ i = 1,2, \dots, N_{triangles}\}$$
$$P2 = \{(x_i^2, y_i^2, z_i^2) | \ i = 1,2, \dots, N_{triangles}\}$$
$$P3 = \{(x_i^3, y_i^3, z_i^3) | \ i = 1,2, \dots, N_{triangles}\}$$

The $i-th$ triangle is denoted by T_i. For each i, it is necessary to label T_i as belonging to a specific section of the satellite. It is chosen to segment out 6 portions of the satellite, thus there are 6 labels for the satellite, including one more label for the background for a total of 7. Labels for each segmented section are: 1) Background, 2) Single Access Antenna, 3) Space Ground Link Antenna, 4) Solar Panels, 5) Supports, 6) Body, and 7) Forward Omni Antenna. (See Fig. 5)

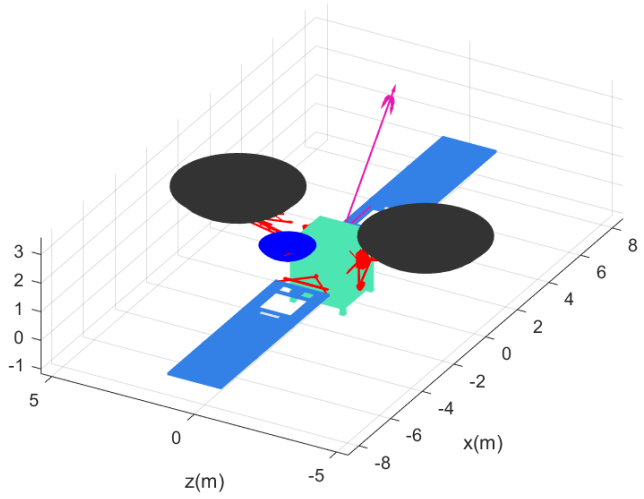


Fig. 5. Segmented CAD model of the TDRS-B Satellite.

Segmented sections are colored as follows: Single Access Antenna (blue), Space Ground Link Antenna (black), 4) Solar Panels (light blue), 5) supports (red), 6) body (teal), 7) Forward omni antenna (magenta).

The labels are necessary for training SalsaNext as the neural network needs to know the answer to what it is trying to segment during the training phase. Lastly, for each i, the normal vector to each of the triangles is collected. These vectors can be computed using a cross product:

$$n_i = \frac{(\langle x_i^3, y_i^3, z_i^3 \rangle - \langle x_i^1, y_i^1, z_i^1 \rangle) \times (\langle x_i^2, y_i^2, z_i^2 \rangle - \langle x_i^1, y_i^1, z_i^1 \rangle)}{\left\| (\langle x_i^3, y_i^3, z_i^3 \rangle - \langle x_i^1, y_i^1, z_i^1 \rangle) \times (\langle x_i^2, y_i^2, z_i^2 \rangle - \langle x_i^1, y_i^1, z_i^1 \rangle) \right\|}$$

The order of the points $(P1, P2, P3)$ ensures that each of the normal vectors are outward pointing and also ensures that the correct side of each triangle is rendered. Direct utilization of the normal vectors does not take place within SalsaNext; rather, these vectors are employed to simulate intensity data, which then serves as input for SalsaNext. A more comprehensive explanation of the intensity data can be found in a subsequent section.

4.2.2 SIMULATION OF POINT CLOUD DATA

Utilizing a labeled CAD model entails capturing snapshots of the model, which are integrated into the training data for SalsaNext. This task is accomplished using a simulated GSFL camera. The objects visible to the camera lie inside of the camera's viewing frustrum. (See Fig. 6)

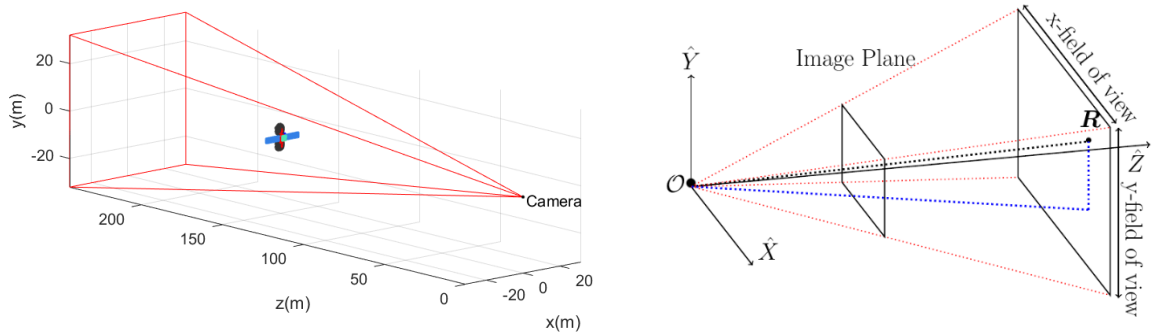


Fig. 6. GSFL field of view frustrum

The GSFL's field of view lies within a frustrum shown in red. The virtual image plane shown is located at f millimeters from the origin $\mathcal{O}$, the range is the distance from $\mathcal{O}$ to R.

The physical camera collects range and intensity data using rays of light (generated via a single laser pulse) that is collected at the camera's 128×128 image plane. When these rays hit an object, the light that bounces off the object is collected by the camera, the time of flight of the light ray (time it takes for the emitted light to return to the camera) is used to determine the 128×128 range values. From the range values, a 3-dimensional point cloud of the object can be computed by using the equations:

$$x = \frac{px \cdot R}{\sqrt{f^2 + px^2 + py^2}}$$
$$y = \frac{py \cdot R}{\sqrt{f^2 + px^2 + py^2}}$$
$$z = \frac{f \cdot R}{\sqrt{f^2 + px^2 + py^2}}$$

Where f is the camera's focal length, R is the range, and

$$\{(px, py) \mid px = -6.35, -6.25, \cdots, 6.25, 6.35, py = -6.35, -6.25, \cdots, 6.25, 6.35\}$$

are 128×128 pixel coordinates (with pixel pitch 0.1mm) with same units as f (millimeters).

In order to simulate the physics of the camera, 128×128 "primary rays" (unit vectors) are generated starting at the GSFL camera's pinhole and pointing in the direction of the center of each of the 16,384 pixels on the image plane. The simulated point cloud pixels are the intersections of the primary rays with the first triangle that it hits (if any). For this, the Möller-Trumbore ray-triangle intersection algorithm [4] is employed. The Möller-Trumbore algorithm returns the (x, y, z) coordinates, and the range to these coordinates. The Möller-Trumbore algorithm is modified slightly so that the normal vectors, and the labels (the label of the triangle that the primary ray hits) are recorded. Gaussian noise is added to the range and the (x, y, z) coordinates to simulate more realistic LiDAR noise.

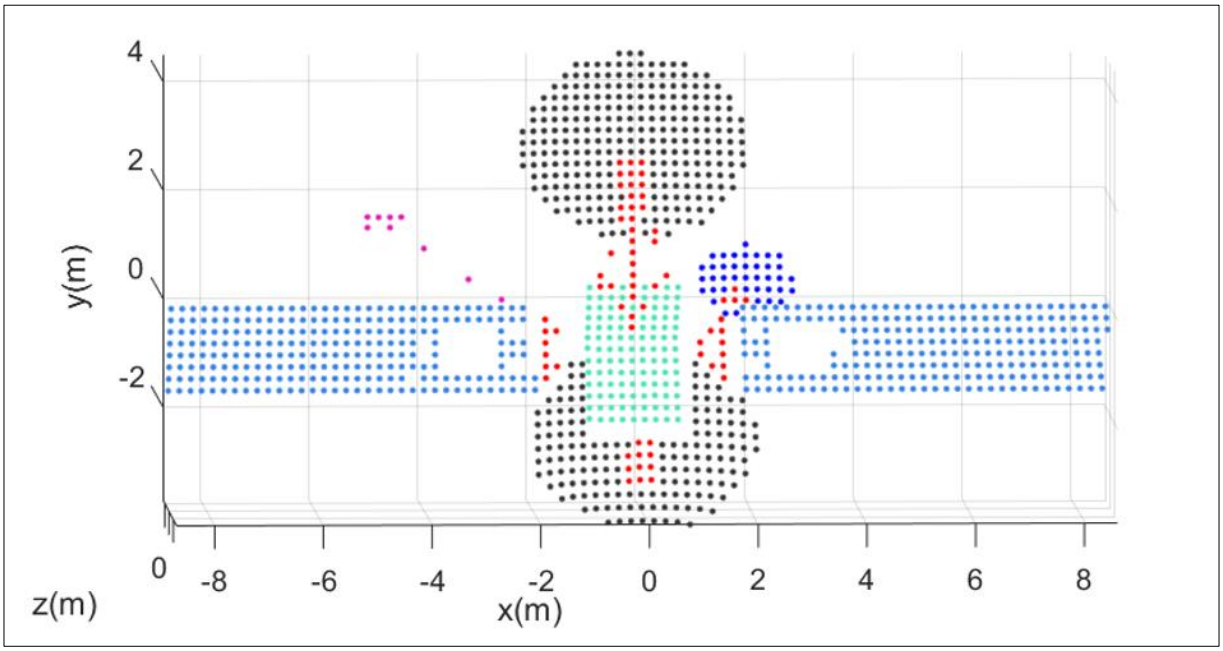
4.2.3 SIMULATION OF INTENSITY COUNTS

The intensity counts are defined to be the number of photons incident on each pixel. This number depends on the range of the target, laser energy, and lens specifications. For simulated intensity counts, this is also dependent on the bidirectional reflectance distribution function (BRDF). The BRDF is a function of the normal vectors, it is widely used in computer graphics and computer vision to simulate how light is reflected at an opaque surface.

The energy reflected on a pixel, in units of photons, is simulated using

$$E_p = \frac{E_{tx}}{R^2 \tan\left(\theta_l/_2\right)} \cdot \rho \cdot \frac{\pi D^2}{4} \cdot IFOV^2 \cdot \eta_{opt} \cdot cos(\alpha).$$

The BRDF used for simulations is the Lambertian model $\rho = \rho_{Lamb}$. The other terms in the above equation are E_{tx}, the total transmitted laser energy (in photons), the range to the target R, the diameter of the aperture D in meters, the instantaneous field of view of the lens $IFOV$ in radians, and η_{opt} is the transmission of the optical system. The angle θ_l is the full angle divergence of the laser in radians and the angle α is the incident angle between the primary rays and normal vectors.



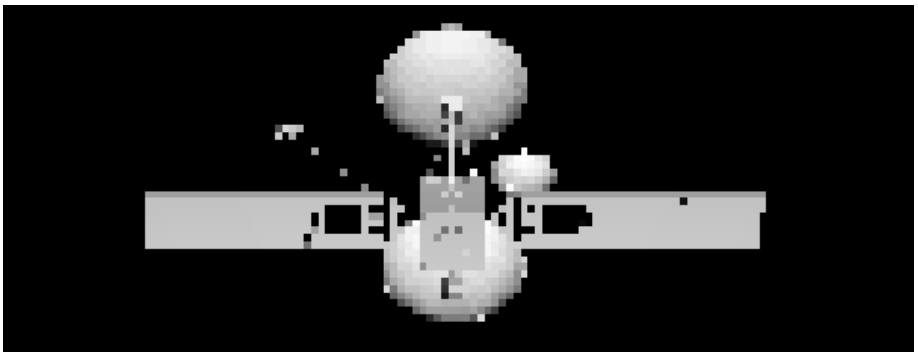


Fig. 7. (Top) Simulated Point cloud. (Bottom) Simulated Intensity.

4.3 SALSANEXT FOR SEMANTIC SEGMENTATION

4.3.1 SALSANEXT DATA PREPARATION

The data is saved in two parts as input and output data. The input data for SalsaNext is a 5-channel organized point cloud (OPC) array which is sized as [128,128,5]. This OPC array is denoted by O. Channel-1 contains x-coordinates, Channel-2 contains y-coordinates, Channel-3 contains z-coordinates, Channel-4 contains Intensity counts, and Channel-5 contains Range values.

The output data is a [128,128] array of integers ranging from 1 to 7 representing the labels. This array is denoted by L. Values in L denote whether the pixel (i, j) is a Background pixel, Single Access Antenna pixel, Space Ground Link Antenna pixel, etc.

To collect 2496 O and L data points:

1) Place the GSFL camera 150 meters from the center of mass of the object, with the camera pointing directly at the center of mass.
2) Use ray tracing and the Lambertian BRDF to generate the 5-Channel array O_i, also generate the Label map L_i for this Organized Point Cloud. Save the data.
3) Rotate the object in a different direction and repeat step 1) and 2) for $i = 1, \cdots 2496$.

The number of images collected does play a factor in the quality of the results, however, it is also important to gather quality images: images of the object at various angles. The simulated camera moves around in a sphere centered at the object and takes pictures of the object.

4.3.2 TRAINING SALSANEXT FOR RSO SEMANTIC SEGMENTATION

SalsaNext attempts to learn a relationship between O_i and i via:

$$L_i = N.N.(O_i),$$

Where $N.N.$ is the SalsaNext DNN. The details of the hidden layers of this DNN can be found in [6]. The data is split into 3 portions, namely the training set, test set, and validation set, having sizes 1498, 499, and 499, respectively. The purpose of the training set is for SalsaNext to learn the geometry of the object. The test set is for testing the model, and the validation set is used to further confirm that a good model has been trained.

4.3.3 SEMANTIC SEGMENTATION RESULTS AND CONCLUSION

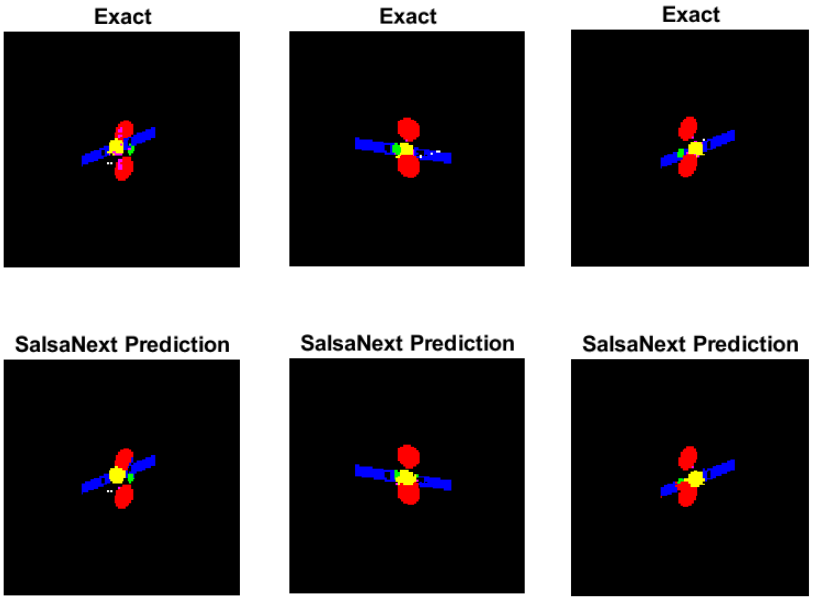


Fig. 8. (Top) Subsets of the Validation data with segmented parts colored differently.
(Bottom) Predicted results from SalsaNext.

The exact solution and the SalsaNext prediction are compared in Fig. 8. The bar graph and confusion matrix that resulted from the validation set are shown in Fig. 9 and Fig. 10 respectively. The data suggests that SalsaNext does an excellent job at differentiating an object and the background. It also does a reasonably decent job at segmenting out substantial portions of the satellite such as the solar panels and the Space Ground Link Antenna. It is difficult for the network to distinguish smaller parts of the antenna such as the supports and the forward omni antenna. The single access antenna was also difficult to segment as it was usually obscured from view. Because SalsaNext was trained with the camera at a specific distance to the object, this model only works well at this distance, however, one can add more images at various distances to the training set to generalize the results to other distances.

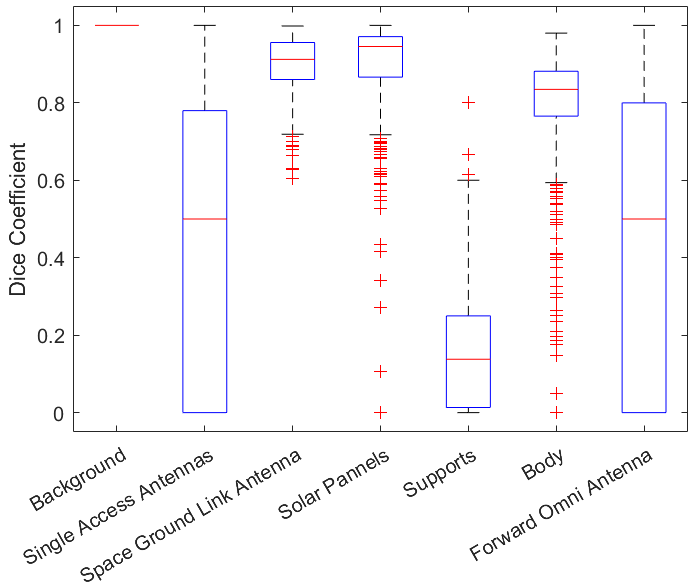


Fig. 9. RSO Dice Accuracy of Test Set

Fig. 10. RSO Semantic Segmentation Confusion Matrix

5. DATA FUSION

5.1 THEORY AND ADVANTAGES

Global Shutter Flash LiDAR has an abundance of data fusion advantages because it provides Organized Point Cloud (OPC) data in real time. The ASC GSFL has a 128-pixel x 128-pixel focal plane array and uses a diffuser to illuminate the entire field of view. This means the camera collects a full frame of organized 3D point cloud data for every laser pulse. This then removes point cloud blurring and distortion, ensuring that everywhere within the field of view is captured in the same pulse. These features of global shutter flash LiDAR are advantageous for easy and accurate data fusion.

As opposed to scanning LiDAR, flash LiDAR and electro-optical (EO) sensors can capture full frames of data simultaneously. This allows for 1-to-1 mapping between pixels in space and time. This concept is illustrated through the next series of figures. Fig. 9 compares the sensor data for a flash LiDAR, a scanning LiDAR [1], and a visible sensor. A flash LiDAR and visible sensor have similar data collection; however, the flash LiDAR has larger pixels. The scanning LiDAR takes points of data along a scanning pattern. In order to register the scanning LiDAR data with a visible sensor frame, the non-uniform scan positions need to be interpolated onto a grid to match the visible sensor sampling locations. This can be a computationally intensive operation. In contrast, the flash LiDAR measures range images in a fixed grid pattern that needs to be correlated with an EO sensor only once.

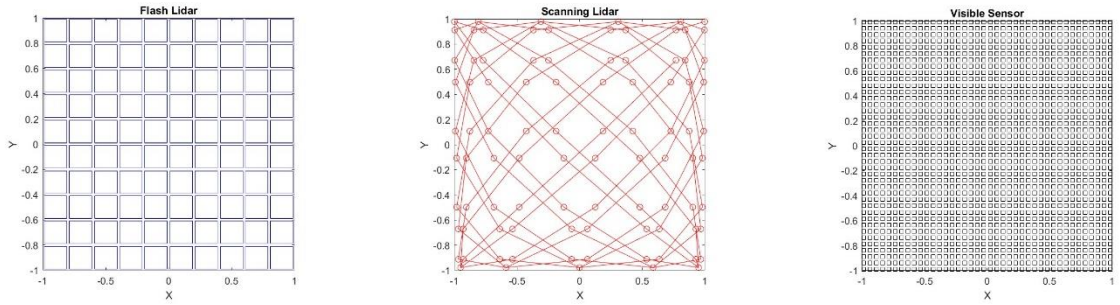


Fig. 9. (Left) Flash LiDAR. (Middle) Scanning LiDAR. (Right) Visible Sensor.

Fig. 10 is a scan pattern comparison between the 3 types of imaging. The flash LiDAR pattern is overlayed on the visible sensor in the first, the scanning pattern is overlayed on the visible sensor in the second, and all three are compared in the third plot. Each pixel on the flash LiDAR sensor can directly map to a group of pixels on the visible sensor.

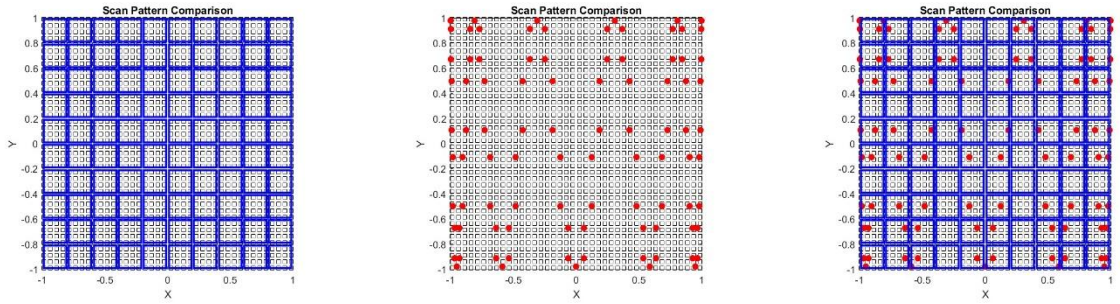


Fig. 10. (Left) Comparison of Flash LiDAR and Visible Sensor. (Middle) Comparison of Scanning LiDAR and Visible Sensor. (Right) Comparison of Flash LiDAR, Scanning LiDAR, and Visible Sensor.

5.2 FLASH LIDAR FUSION WITH VISIBLE AND IR DATA

ASC has begun preliminary work on data fusion between the GSFL sensor and both visible and long wave infrared (LWIR) data. The GSFL inherently measures point cloud data as an organized point cloud array. Each pixel in the FPA has a vector with corresponding data (X, Y, Z, Range, Intensity). It is therefore straightforward to fuse GSFL data with other types of camera data. The array can be expanded to include more information (X, Y, Z, Range, Intensity, R, G, B, temperature, etc.) in the vector for each pixel in each frame.

To demonstrate this concept, footage of a parking lot with palm trees and bushes was taken with a GSFL, a visible camera, and an IR camera. Fig. 11 shows the results: colorized, organized 3D point clouds. On the left, the GSFL 3D point cloud data is fused with RGB data from a visible camera, while on the right the same data is fused with temperature data from a long wave infrared (LWIR) camera. This results in a seven-dimensional dataset that can convey detailed information about a three-dimensional scene. Since all of the sensors in this suite are arrayed detectors, the registration overhead is minimal allowing for more computational resources to be devoted to intelligent analysis of the measured data.

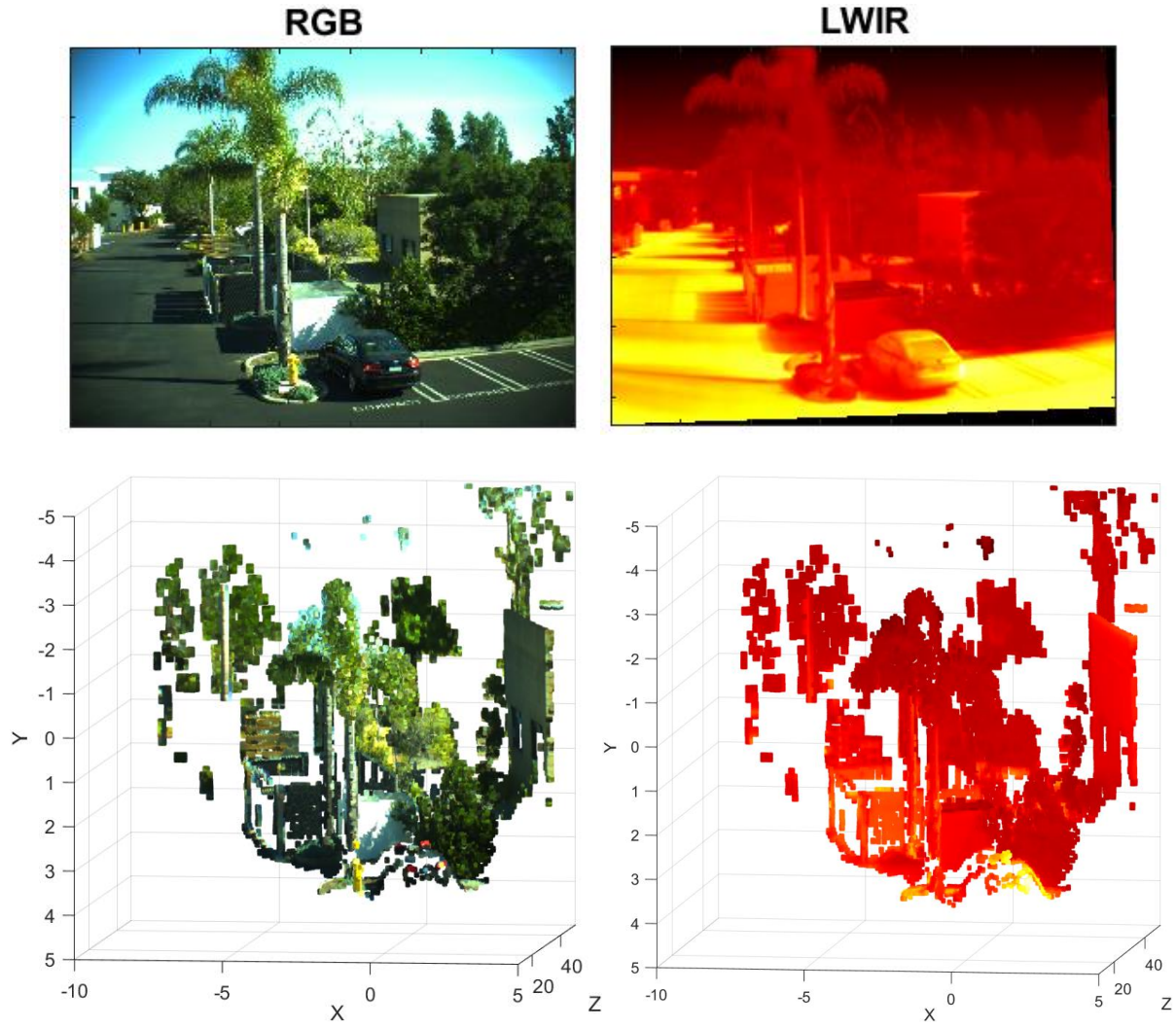


Fig. 11. (Top Left) RGB Image. (Top Right) LWIR Image. (Bottom Left) RBG image data fused with GSFL organized point cloud of scene. (Bottom Right) LWIR image data fused with GSFL organized point cloud of scene.

6. CONCLUSION AND FUTURE WORK

ASC is continuously working to extend both software and hardware configurability. The focus of ongoing work at ASC is the implementation of ML/AI algorithms. Regarding Semantic Segmentation, SalsaNext undergoes training using densely populated point clouds, employing a designated focal length. The next step is to check if SalsaNext can segment objects captured using different focal lengths by inputting a rescaled version of the point cloud into SalsaNext (scaled so that it matches the focal length of the highly dense point cloud). ASC is also working to refine the mechanical design concept within practical SWaP budgets.

ASC is actively engaged in ongoing research and practical implementation of data fusion involving Global Shutter Flash LiDAR and various other data types. The goal of data fusion research is to create systems that can communicate multiple different classes of useful information about a scene. Data fusion can also support and improve semantic segmentation and other ML/AI algorithms in future work.

7. REFERENCES

[1] Junya Wang, Gaofei Zhang, and Zheng You. Design rules for dense and rapid Lissajous scanning, *Microsystems & Nanoengineering* (2020) 6:101.
[2] Lane Fuller, Amy Carl, Joe Spagnolia, Brad Short, Michael Dahlin. Photon counting linear mode global shutter flash LIDAR for improved range performance, Proc. SPIE 12110, Laser Radar Technology and Applications XXVII, 1211005 (3 June 2022); https://doi.org/10.1117/12.2619047
[3] Lane Fuller, Robert Karl, Jr., Bruce Anderson, and Max Lee-Roller. Development of a versatile LiDAR point cloud simulation testbed for advanced RSO algorithms, Proc. The Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference (September 2022).
[4] Möller, T. T. (1997). Fast, Minimum Storage Ray-Triangle Intersection. *Journal of Graphics Tools. 2*, 21-28.
[5] NASA. (2023, August 15). *3D Models*. Retrieved from NASA 3D Resources: https://nasa3d.arc.nasa.gov/models
[6] Tiago Cortinhal, G. T. (2020). SalsaNext: Fast, Uncertainty-aware Semantic Segmentation of LiDAR point clouds for autonomous driving, *Advances in Visual Computing*, 207-222.