Super-Resolution Object Characterization in Low Earth Orbit (SROC LEO)

Stacey Jones PhD | DSc
O Analytics Incorporated

ABSTRACT

Dependence on the Space Surveillance Network (SSN) or other monitoring networks that are comprised primarily of traditional sensing technologies complicates real-time object detection and classification. The quality of imagery emanating from these sources may be adversely affected by sensor age, physical location and/or orientation. Moreover, status quo artificial intelligence/machine learning (AI/ML) solutions alone are generally not well suited for high performance on blurred and otherwise degraded imagery, as they generally don't bode well with obscure representations. We present a Super-Resolved Object Characterization in Low Earth Orbit (SROC LEO) approach that optimally exploits blurred imagery consistent with that made available from a significant number of SSN legacy sensors employing optical technology. By uniquely combining and adapting proven fast kinematic and super-resolution methods, early indications are that our SROC LEO concept will surpass performance of status-quo satellite classification to achieve improvement by up to a factor of four (4) in each of the two (2) pixel plane dimensions.

1.0 INTRODUCTION

The application of super-resolution techniques to satellite-to-satellite imagery, and investigation of the effects of these approaches on object detection and characterization performance is scarce. Such interplay between super-resolution techniques and object detection frameworks is vastly enigmatic and consequently remains largely unexplored [1]. Notwithstanding, there is work in remote sensing Earth imagery, medical imaging and facial recognition from which a basis for super-resolution of non-Earth imagery (NEI) can be established. Examples include Landsat [2], Ground Video [3][4], UAS photogrammetry [5], Optical Coherence Tomography (OCT) [6], and facial recognition [7].

Detection and conclusive identification of satellites in an increasingly congested low Earth orbit (LEO) is complex, and while improving at an impressive rate, remains limited. This limitation impairs the ability to protect assets in space from bad or irresponsible actors and impedes ensuring in-space operations are free from interference or obstruction. For commercial operators, definitive object characterization in an operating region remediates damage of satellite constellations due to collision. In both military and industry scenarios, risk abatement starts with adroit situational and domain awareness.

Dependence on the Satellite Surveillance Network (SSN) or other monitoring networks that are comprised primarily of traditional sensing technologies and infrastructure further limits object classification. Moreover, status quo artificial intelligence/machine learning (AI/ML) solutions alone may not be well suited for this non-native application as they generally don't bode well in noisy or otherwise visually complicated environments such as that existing in low Earth orbit (LEO).

Visually complicated environments with respect to an object in LEO can take on many forms. Some common forms include (1) objects that are a far range, (2) objects that are at far range with competing background, (3) partial objects in the field of view (FOV), (4) partial objects at very close range and various combinations thereof.

(1) Objects at Far Range
Objects that are at far range introduce complexities primarily associated with low pixel count occupation. These scenarios often lack distinction that is difficult to resolve and therefore a challenge to the classification processing. Relative lighting (i.e., dim object versus background) may also be a significant factor.

Fig. 1 Example Scenario – Object at Far Range (20OS2)

(2) Objects at Far Range with Competing Background

Objects that are at far range with competing backgrounds introduce two-fold complexities associated with low pixel count occupation and visually noisy areas in the FOV. These scenarios not only lack distinction difficult to resolve as it relates to the object of interest, but the object itself may appear to be indistinguishable from the noisy areas. Noisy areas may be of a variety of types such as lighting or contrast differences, weather, Earth perspective, or debris in the FOV.

Fig. 2 Example Scenario – Object at Far Range with Background (20DS3)

(3) Partial Objects within the FOV

Objects that are not fully encapsulated in the FOV or frame do not provide the best opportunity for accurate classification as there is only partial representation available for processing. These scenarios may be a result of the object just entering or beginning to vacate the FOV.

Fig. 3 Example Scenario – Partial Objects within the FOV (20DS15)

(4) Objects at Close range with Competing Background
Objects that are at closer ranges occupy more of the full pixel plane. Moreover, those with textured backgrounds introduce processing that may not lend well to some timesaving filtering or predictive techniques. While the computational effort to conclusive classify these scenarios may be tedious, they are especially relevant when considering In-Space Servicing, Assembly and Manufacturing (ISAM) activity such as Rendezvous Proximity Operations (RPO), as well as Space Situational Awareness (SSA) efforts. Uneven or inconsistent illumination further complicates scenarios in this broad category.

Fig. 4 Example Scenario – Object at Close Range with Textured Background (20DS13)

There are of course countless variations as well as additional scenarios that can occur in low Earth orbit (LEO) with respect to objects. In this paper and the associated work, we examined a limited set of realistic scenarios as a point of departure for performing high fidelity object characterization under ideal and imperfect conditions.

More specifically, we explore fast kinematic and machine learning approaches that with a specified order and a unique combination of adapted techniques, adds value in the present given the existing optical sensors currently deployed in LEO to capture non-Earth imagery (NEI). Moreover, our resulting Super-Resolution Object Characterization in Low Earth Orbit (SROC LEO) accommodates continuous improvement that positions it well for further contribution to object characterization as NEI optical sensing technologies advance in terms of image quality.

2.0 SROC LEO OVERVIEW

Super-Resolution Object Characterization in Low Earth Orbit (SROC LEO) is a deep learning adaptation and application of principles from a medical imaging solution to discover potential biomarkers (i.e., for Alzheimer's diagnosis) using super-resolution [6] and facial recognition pose estimation concepts, combined with an innovative target chipping method specifically tailored for satellite pixel region determination. The purpose of SROC LEO is to significantly improve near real-time space object characterization for low Earth orbit (LEO) targets – most specifically satellites. SROC LEO builds upon, extends and modifies an existing GAN architecture proven to identify biomarkers and faces, incorporates novel and efficient chipping preprocessing, and includes several computational and processing adaptations from related deep learning concepts to yield superior space object characterization. Applying principles of an Optical Coherence Tomography (OCT) retinal Generative Adversarial Network (GAN) architecture [8][9] and pixel-plane image processing techniques in a prescribed order indicated promising measurable and visibly impressive performance with respect to the endgame of satellite classification.

SROC LEO deep satellite image detection, super-resolution, and classification consists of three core modules. The satellite detection module detects and extracts the region of interest (ROI) referred to as a target (i.e., satellite) chip using novel algorithms that include edge detection (i.e., Innovative Target Chipping (ITC)). These satellite chips undergo an adapted super-resolution series by an up-scale factor of 8/16 using a further adapted Multi-scale (MS) SRGAN super-resolution module followed by a Resnet-18 satellite classifier to initially identify the type of the satellite in the region of interest. The multi-stage super-resolution network is jointly trained with the satellite classifier. The final classification step involves correlating features extracted during the process flow. The performance of SROC LEO has intrinsic dependencies between its core modules.

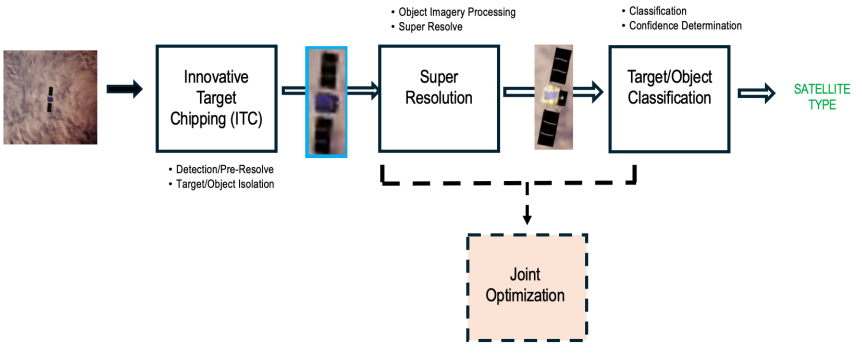

Fig. 5 SROC LEO Block Diagram w/core modules

It is important to note that the datasets used for our investigation were synthetic imagery (i.e., computer-generated imagery (CGI)). This approach was both out of necessity and convenience. The amount of real non-Earth imagery (NEI) available through public and relational sources were and remain limited for this level of deep learning investigation that can call for tens of thousands of images or more. As such, several techniques were used to create our full dataset using a variety of augmentation and degradation approaches – a number of which also employed several core principles of GANs [10][11][12]. However, a benefit and hence intrinsic convenience of the use of CGI was the ability to more openly distribute our concept and multiple stages of results for comment and review without regard to or concern with disclosing possible sensitive information.

3.0 INNOVATIVE TARGET CHIPPING (ITC)

The Innovative Target Chipping (ITC) applies a novel approach to detect the presence of anthropogenic objects (i.e., satellites, rocket bodies, debris) from single JPEG images and provides the pixel coordinates that define and encapsulate the detected object. The determined vertices that constitute a bounding box are provided as inputs to the subsequent multi-layered super-resolution enhancement process. The ITC construction of a layered edge map that is modified using image gradient and color statistical properties are important aspects of the ITC.

Within the ITC, an original edge map is created using an enhanced 3 x 3 kernel Sobel [13] convolution. This approach emphasizes edges [14] such that the edge map modifiers were based on the hypothesis that anthropomorphic objects would intrinsically have anisotropic spatial gradient distributions of the variance and standard deviations. To avoid interference of each frame image boundary with respect to candidate objects in the field of view (FOV), all gradient angles converted to a numeric modulus were applied multiplicatively to the edge map metrics. Due to the wide range of scenarios, high luminance cases occurred due to natural reflection from clouds and daytime blue atmosphere and with stark contrast that included dark space background. In addition to using histogram-like distributions for association of weighted segments, proximity functions were also used to identify dominant color via inverse Euclidean distance in pixel space [15]. The resulting joint constraint threshold detection model allowed for proper detection of the object of interest for a plethora of evaluated scenarios. The pixel coordinates assigned to the bounding box vertices were based on height and width traversal of the confidence-based decision thresholds.

Examples of results of the ITC satellite detection module with extraction of the region of interest (ROI) referred to as target (i.e., satellite) chips for aforementioned scenario types are presented in Fig. 6-10.

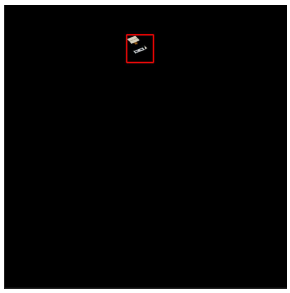

Fig. 6 Example 1 ITC ROI extraction (20DS06)

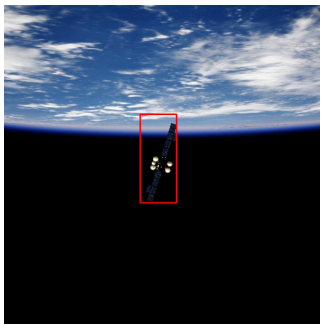

Fig. 7 Example 2 ITC ROI extraction (20DS19)

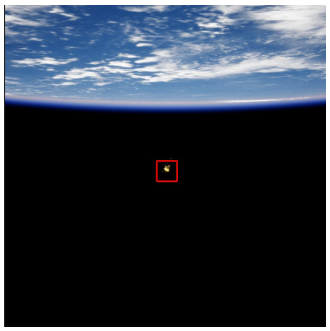

Fig. 8 Example 3 ITC ROI extraction (20DS03)

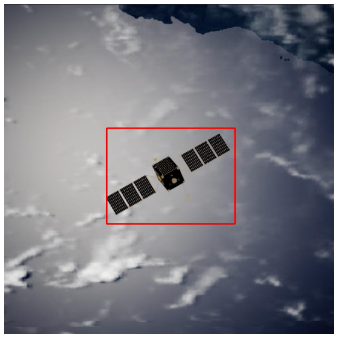

Fig. 9 Example 4 ITC ROI extraction (20DS04)

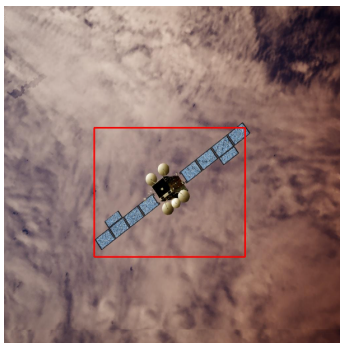

Fig. 10 Example 5 ITC ROI extraction (20DS17)

Fig. 11 Example 6 ITC ROI extraction (20DS14)

Fig. 12 Example 7 ITC ROI extraction (20DS15)

4.0 SUPER-RESOLUTION

The primary goal of super-resolution is to enhance the pixelation and quality of low-resolution images that may otherwise be difficult to discern in terms of their content. Our interest in the application of super-resolution specifically relates to improving visual representation and ultimately achieving better classification of space objects within NEI by autonomous systems. In simplest terms, our work focuses on addressing blurred, noisy or otherwise distorted aspects of an image such that objects of interest within the FOV appear and are in fact measurably clearer such that our confidence in properly classifying said objects approaches 100%. Super-resolution facilitates the extraction of finer details and characteristics from training with many images, which is crucial for accurate classification. It's integration into the overall SROC LEO process flow aims to enhance the overall accuracy and robustness of the classifier, particularly in scenarios where small target chips are low resolution making it exceptionally difficult for accurate classification.

The SROC LEO super resolution functionality is achieved via an adapted Multistage Super Resolution Generative Adversarial Network (MSRGAN) architecture. The resulting MSRGAN has two configuration types, for achieving 8x and 16x super resolution respectively. These models use multiple discriminators in the different resolution configurations to guide the super resolution process of the generator. Theoretically, the generator and the discriminator are required to learn and grow together at a similar pace so that the generator can flawlessly match the distribution of data. However early on in several cases the generator tended to be left behind without useful gradients and ceased improving the quality of the synthesized images. We alleviated this and related issues by incorporating multiple discriminators in multiple stages and arrived at a satisfactory point with the architectures depicted in Fig.13 and Fig.14.

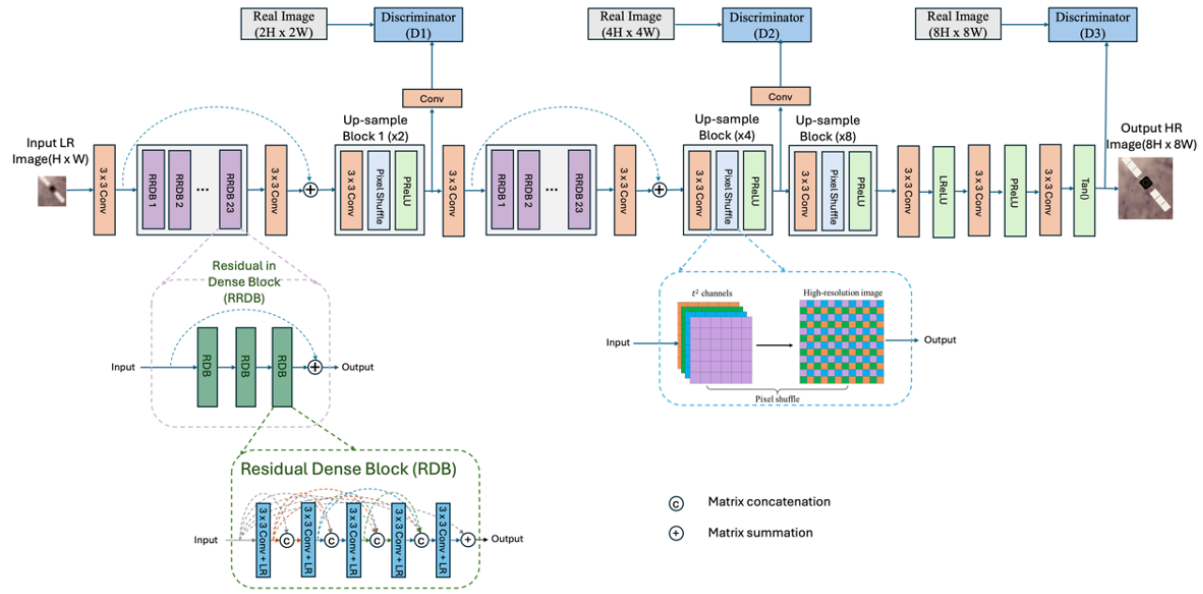

Fig. 13 SROC LEO Generative Architecture (8x)

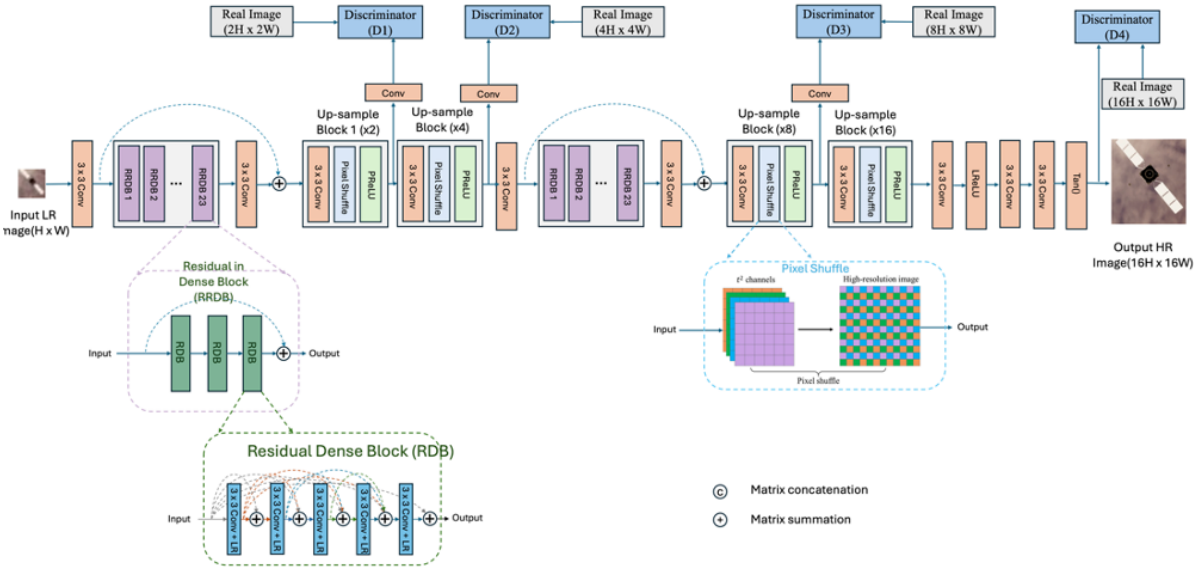

Fig. 14 SROC LEO Generative Architecture (16x)

Background and Select Details

Generator Architecture: The Residual-in-Residual Dense Block (RRDB) used for the generator architecture interconnects all layers within a residual block to boost the network's capacity [16]. Residual scaling involves multiplying a parameter β within the range [0, 1] to the residual, preventing instability [17]. For a ×16 upscaling factor, the network's first half executes ×4 upsampling (using two successive PixelShuffle (×2)), while the latter half completes the remaining ×4 upsampling. Similarly, in the case of ×8 upsampling factor, the initial half performs ×2 upsampling, and the latter half handles the remaining ×4 upsampling. Employing a multi-stage generator yields outputs from various intermediary layers, not just the final one [18].

Discriminators Architecture: The main purpose of the discriminator is to classify between the generated outputs and the ground truths at each intermediary stages. In this SROC LEO work, multiple discriminators are being used to

enhance the discriminatory aspects of the GAN. Each of these discriminators is a PatchGAN [19] in which a convolutional neural network classifies an image as fake or real by focusing on penalizing it at the scale of local image patches of size N × N. Patch-based discriminators ($D_i \in \{i=1, 2, 3, 4\}$) progressively contribute to learning finer details (higher resolution) in every scale. Stage-specific patch-based discriminators are employed at multiple SR stages to further refine generated images' texture towards indistinguishability from I_{HR}. Addressing critical concerns in GAN training, spectral normalization (SN) is applied in each discriminator convolutional layer to stabilize adversarial training against exploding and vanishing gradients [10][20]. The former causes training instability, while the latter results in either bad local minima or training stalling before convergence [10]. Additionally, each discriminator after several convolutional layers has two fully connected layers, followed by batch normalization and leaky ReLU activation function. At the top of these hidden layers lies a single neuron with a Sigmoid activation function.

PixelShuffle Layer: The purpose of PixelShuffle is to increase the resolution of the feature maps. This method takes a feature map as input, which consists of spatial information (width and height) along with a certain depth (number of channels). Through PixelShuffle(x2), the elements within each spatial neighborhood of the feature map are rearranged, resulting in an effective upscaling of the spatial resolution. Specifically, the spatial resolution is doubled (increased by a factor of 2) while the depth is reduced by a factor equal to the square of the upscaling factor, ensuring that information from adjacent channels is effectively combined. This process enables the model to enhance image resolution without losing valuable information, making it particularly useful in tasks requiring high-quality image reconstruction or enhancement [21].

Loss Functions: We used adversarial, L1 and perceptual in our SROC LEO work thus far and at the time of this writing concluded that our investigation was not sufficiently complete to include details here.

5.0 CLASSIFICATION

Within the context of SROC LEO, recall that our goal is to classify widely varying satellite target chips in terms of scenarios and size. Our assessment of required robustness of a classifier was heavily influenced by the ability to handle widely varying and unpredictable target chip sizes resulting from the ITC preprocessing. We determined that given its intrinsic robustness and efficiency when applied to very deep and very large networks, that employing a ResNet [22] classifier would be sufficient.

Background: Microsoft Research Group introduced the ResNet [22] architecture to build a deeper neural network and to address the issue of vanishing and exploding gradients. A key innovation of this network is the skip connection which helps to backpropagate loss to the initial layer. By utilizing the skip connection, the ResNet architecture achieved state-of-the-art classification and object detection performance in computer vision benchmark datasets. In ResNet, skip connections, also known as shortcut connections, are fundamental to the architecture's design. They enable the network to bypass one or more layers, allowing the gradient to flow more directly during training. The simplest form of skip connection in ResNet is identity mapping where the input to a layer is directly added to its output. Here the output of a ResNet layer is the sum of its input and the output of the layer's convolutional operations. This skip connection typically bypasses one or more convolutional layers.

In our SROC LEO classification approach, through reasonable adaptation, we utilized a shallower version of ResNet-18, without any stride. This variant of the Residual Network (ResNet-18) architecture was specifically developed with 18 layers. It consists of basic building blocks called residual blocks. An overview of the architecture of the SROC LEO satellite classifier merged with general background is as follows:

- *Input Image*: The input to a Shallower ResNet is typically an RGB image with a variable size, such as 10x20, 30x12,120x180 pixels.
- *Initial Convolutional Layer*: The input image passes through an initial convolutional layer with 64 filters (kernels) of size 3x3, with a stride of 1. The convolution is followed by a batch normalization and a ReLU activation function. The output feature maps have the same spatial dimensions due to the stride set to one.
- *Residual Blocks*: SROC LEO ResNet consists of four residual blocks, each containing multiple convolutional layers. Within each residual block, there are two convolutional layers, except for the first block. After each convolutional layer, there is a batch normalization and a ReLU activation function. The first convolutional layer within each block has a filter size of 3x3, and the number of filters gradually increases from 64 to 128 across the blocks. Each convolutional layer in every block utilizes a stride of 1. Therefore, the spatial size of feature

maps remains the same. Shortcut connections (i.e., skip connections) are used to add the input of a block to the output of the block.

Global Average Pooling (GAP): After the last residual block, an adaptive global average pooling layer is applied. After global average pooling on the feature map to contain rigorous spatial information. This layer averages the spatial dimensions of each feature map independently, resulting in a fixed-size feature vector regardless of the input size.

Fully Connected Layer (Output Layer): Finally, a fully connected layer with a SoftMax activation function produces the output logits. The number of neurons in the output layer corresponds to the number of classes in the classification task.

Output: The output of the fully connected layer represents the predicted probabilities for each class.

Overall, our classifier is a relatively shallow variant of the ResNet architecture, designed to balance computational efficiency with performance. Despite its shallow structure, it has been shown to achieve promising performance on satellite image classification tasks.

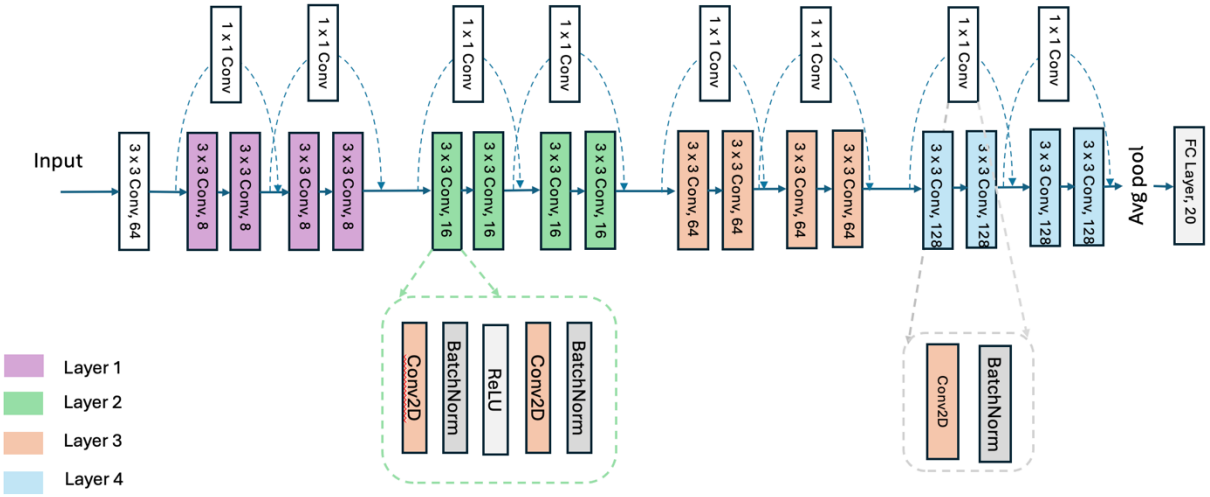

Fig. 15 SROC LEO ResNet-18 Classifier architecture

6.0 TRAINING

As was the case with the SROC LEO super-resolution loss function, at the time of this writing we concluded that our investigation of candidate training approaches was not sufficiently complete to include a full summary of details here. However, our initial training details are presented here for both the super-resolution and classifier, and our joint optimization implementation.

Super-resolution Training Initial (Example Run)

There were 20 classes of satellite images with varying backgrounds. We used the ITC module to detect and crop the target chips. Then we downsampled applying the wavelet decomposition technique. After that we split the data set into two sets. Working with a 10,000 image data set, classes were partitioned as training versus test sets (i.e., 75% | 25 %) to ensure that there is no overlapping of the training and test data for generalization. We used the Adam optimizer to train the super-resolution algorithm with a learning rate of 0.00025 and 0.0001 for the generator and discriminator, respectively. The proposed framework benefits from the relative influence of different loss functions, which is guided by the regularization parameters: $\lambda_{GAN} = 0.01$, $\lambda_{l1} = 1.0$, $\lambda_{lpips} = 0.001$, $\alpha_1 = 0.001$, $\alpha_2 = 0.005$, $\alpha_3 = 0.01$, and $\alpha_4 = 1$. For qualitative comparisons, results were reported for both upscaling factors of ×8 and ×16. The experiment is performed using the PyTorch framework with two NVIDIA Titan-RTX GPUs.

Results on Initial Test set:

Fig. 16 and Fig. 17 show the generator and discriminator training loss curves over time, demonstrating that the generator loss decreases as epochs progress. The MSR-GAN generator was initialized from scratch. Similarly, Fig.18 and Fig. 19 present the training loss curves for the 16x MSGAN. For 8x downsampled images, the average PSNR and

SSIM were 23.52 dB and 0.606, respectively. For 16x downsampled images, the values were 21.26 dB and 0.606. The computational times were 0.1364 s/image for 8x and 0.096 s/image for 16x, using an NVIDIA Titan-RTX GPU.

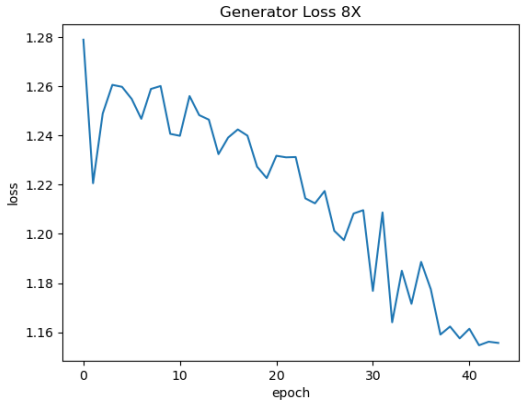

Fig. 16 Training loss curve of the generator of 8x super resolution by MSR-GAN.

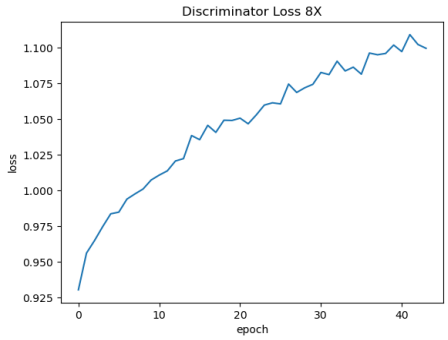

Fig. 17 Training loss curve of the discriminator of 8x super-resolution by MSR-GAN.

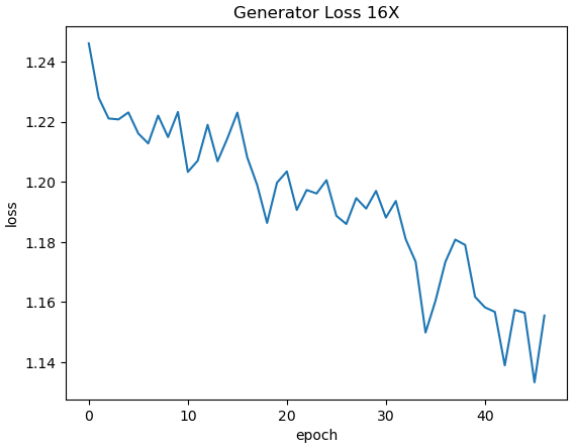

Fig. 18: Training loss curve of the generator of 16x super-resolution by MSR-GAN

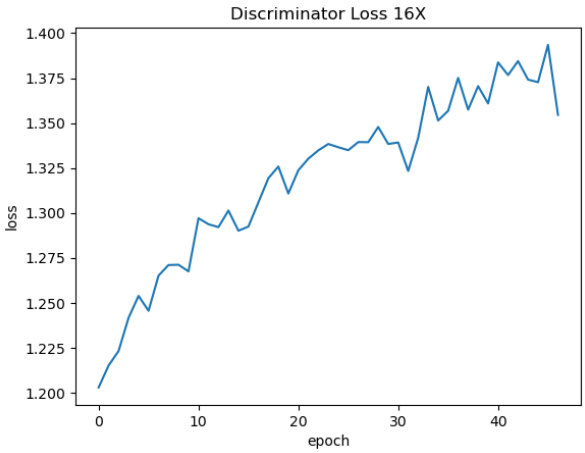

Fig. 19: Training loss curve of the discriminator of 16x super-resolution by MSR-GAN

Table 1: The average PSNR and SSIM of multi-stage super-resolution network.

	PSNR [dB]	SSIM	Computational Time (s)
16x	21.26	0.695	0.096
8x	23.52	0.606	0.1364

Classifier Preliminary Training (Example Run):
Again, using the satellite dataset consists of 20 classes (each class consists of 500 samples), we divided the training, testing, and validation at an 80:15:5 ratio. We used the Adam optimizer and a learning rate of 0.001 to train the shallow ResNet classifier. After each epoch, the learning rate decreased exponentially by 0.95. We trained the classifier for 92 epochs and plotted the validation accuracy curve in Fig. 21. We used the categorical cross-entropy cost function to calculate the loss. Finally, we finetuned the classifier for 8 epochs for variable ITC target chips, for this regard, we use a learning rate of 0.0001, which decreased exponentially with a factor of 0.9 in elapsed of each epoch. During training the ITC target chips, we use the original resolution and super-resolution target chips. The distribution of training, testing, and validation sets is similar to that described before.

Results on Validation Dataset
Fig. 21 depicts the classifier's validation set performance (accuracy) with the elapsed number of epochs. From there, we can see that the classifier's performance became saturated after 20 epochs. Fig. 20 and Fig. 22 illustrate the classifier's training and validation loss curves, respectively. From there, we can see loss become constant after ten epochs. However, after 70 epochs, classifier loss starts to increase due to overfitting. The classifier is trained on high-resolution images and tested on i) original resolution, ii) 8X downsampled, and iii) 16X downsampled images. The confusion matrix of the test set for three scenarios is depicted in Fig. 23 a) – c). The average accuracy of the original resolution, 8X downsampled, and 16X downsampled test set is 92.12%, 43.07%, and 18.42%, respectively. From there, we can see the classification performance is dropped due to the downsampling of the test images. To improve the performance in down-sampled images, we trained the target chips using all resolutions (down-sampled, super-resolved, and original). The confusion matrix for the all-resolution test sets and the performance of each class are reported in Fig. 24. From there, we can see that the average accuracy on the test set is 97.4%. We also observed that every class performance is more than 90% except for class-4, class-8, and class-12. From Fig. 24, we also found that the classification performance substantially improves after super-resolving the degraded target chips. The average elapsed time for each target classification is 2.5 milliseconds on the NVIDIA RTX 6000 Ada GPU system.

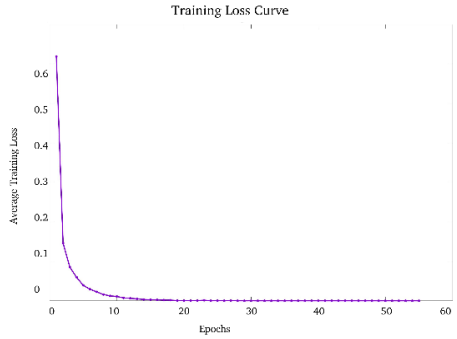

Fig. 20: Training loss curve of the classifier.

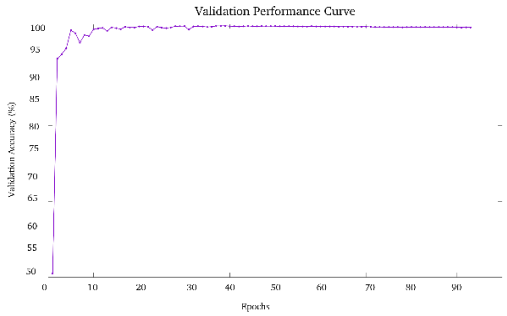

Fig. 21: Performance of the classifier on the validation dataset.

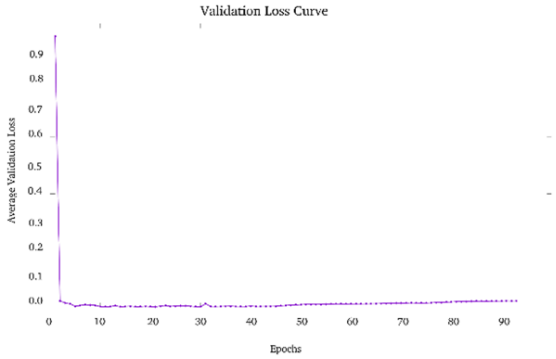

Fig. 22: Validation loss curve of the classifier

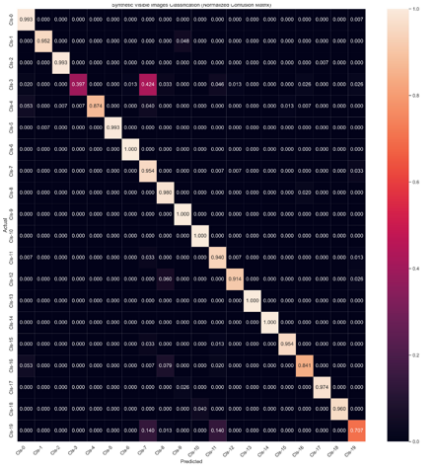

Fig. 23 a) Test on original resolution images

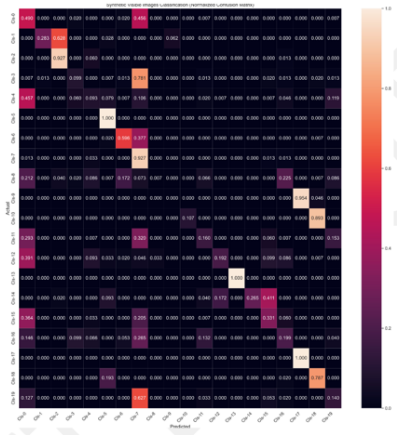

Fig. 23 b) Test on downsampled (8X) images

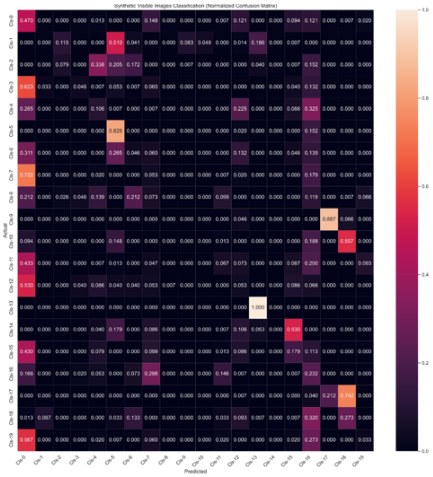

Fig. 23 c) Test on downsampled (16X) images

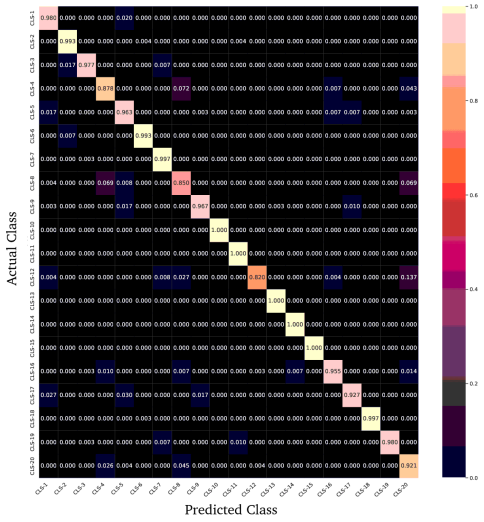

Figure 24 Confusion Matrices of the test set on all-resolution target chips. The classifier is trained using all-resolution (original, down-sampled and super-resolution) satellite images.

In summary, we utilized an SROC LEO ResNet-18 classifier with stride one for variable target classification. We analyzed the training and validation loss curves to obtain the optimal checkpoint. We found that training the classifier on original target chips and testing on downsampled target chips degrades performance. To address this issue, we used super-resolution on the downsampled target chips and trained the classifier using original, downsampled, and super-resolved target chips. By taking these steps, we achieved better performance on degraded target chips, as illustrated in Fig. 24.

7.0 CONCLUSION AND DISCUSSION

This paper presents very promising Super Resolution Object Characterization in low Earth orbit (SROC LEO) results in a newly investigated and complex area of detection and conclusive identification of satellites in an increasingly congested LEO using non-Earth imagery (NEI). At the time of this writing, our SROC LEO loss and training approaches were still in process and consequently were covered with limited scope herein. Notwithstanding, we were able to establish a reliable foundation with strong technical merit for subsequent exploration. To this end we built and deployed initial SROC LEO demonstration prototype versions on two distinct platforms – one of which is of pre-space flight technology readiness level. This provides a repeatable environment for subsequent SROC LEO advancement and improvement.

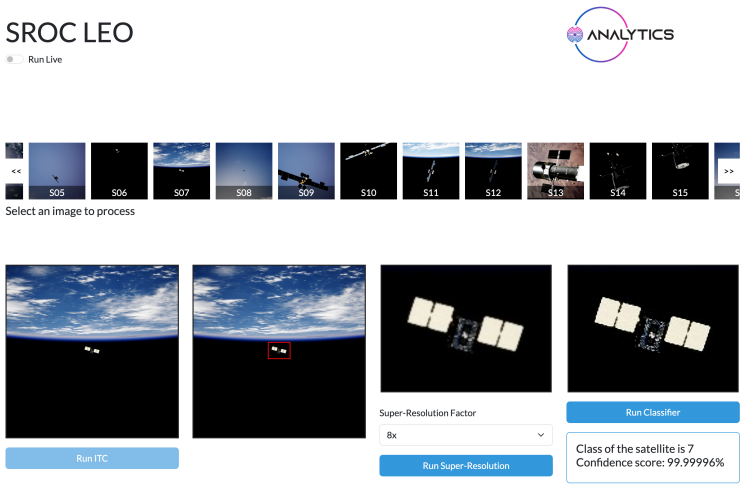

Fig. 25 SROC LEO Demonstration Prototype (Example Screenshot)

ACKNOWLEDGEMENT

The author wishes to thank the Department of the Air Force (DAF) Air Force Research Laboratory Space Vehicles Directorate (RVSV) engineer Matthew Cleal for his technical commentary and feedback from a Space Force perspective. Additionally, the author thanks the research team from the West Virginia University Deep Learning Laboratory (WVU DLL) comprised of faculty and four (4) doctoral students, for their multifaceted contribution to this effort. Lastly, the author thanks the O Analytics research and development (R&D) and concept of operations (CONOPS) team members who helped to shape and demonstrate the strength and versatility of this effort.

REFERENCES

[1] Shermeyer, J., & Van Etten, A. (2019). The effects of super-resolution on object detection performance in satellite imagery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 0-0).

[2] Pouliot, D., Latifovic, R., Pasher, J., and Duffe, J. (2018). Landsat Super-Resolution Enhancement Using Convolution Neural Networks and Sentinel-2 for Training. *Remote Sensing*, 10(3):394, Mar. 2018.

[3] Luo, Y., Zhou, L., Wang, S. and Wang, Z. (2017). Video Satellite Imagery Super Resolution via Convolutional Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, 14(12):2398–2402, Dec. 2017.

[4] Xiao A., Wang Z., Wang L., Ren Y. (2018). Super-Resolution for “Jilin-1” Satellite Video Imagery via a Convolutional Network. *Sensors*; 18(4):1194.

[5] Pashaei M, Starek MJ, Kamangir H, Berryhill J. (2020). Deep Learning-Based Single Image Super-Resolution: An Investigation for Dense Scene Reconstruction with UAS Photogrammetry. *Remote Sensing*; 12(11):1757.

[6] Jeihouni, P, Dehzangi, O., Amireskandari, A., Rezai, A. and Nasrabadi, N. (2021). MultiSDGAN: Translation of OCT Images to Super resolved Segmentation Labels Using Multi-Discriminators in Multi-Stages. *IEEE Journal of Biomedical and Health Informatics*. PP. 1-1. 10.1109/JBHI.2021.3110265.

[7] Jiang, J., Wang, C., Liu, X. and Ma, J. (2023). Deep learning-based face super resolution: A survey, *ACM Comput. Surv.*, vol. 55, no. 1, pp. 1-36, Jan. 2023.

[8] Kugelman, J., Alonso-Caneiro, D., Read, S. and Collins, M. (2022). A review of generative adversarial network applications in optical coherence tomography image analysis. *Journal of optometry*. 15. 10.1016/j.optom.2022.09.004.

[9] Zheng C, Xie X, Zhou K, Chen B, Chen J, Ye H, Li W, Qiao T, Gao S, Yang J, Liu J. (2020) Assessment of Generative Adversarial Networks Model for Synthetic Optical Coherence Tomography Images of Retinal Disorders. *Transl Vis Sci Technol.* 2020 May 27;9(2):29. doi: 10.1167/tvst.9.2.29. PMID: 32832202; PMCID: PMC7410116.

[10] Zhang, K., Liang, J., Van Gool, L., & Timofte, R.(2021). Designing a practical degradation model for deep blind image super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4791-4800).

[11] Cubuk, E. D., Zoph, B., Shlens, J., & Le, Q. V. (2020). Randaugment: Practical automated data augmentation with a reduced search space. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 702-703)

[12] Bulat, A., Yang, J., & Tzimiropoulos, G. (2018). To learn image super-resolution, use a gan to learn how to do image degradation first. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 185-200).

[13] Sobel, I. (1990). An isotropic 3× 3 gradient operator, machine vision for three–dimensional scenes. *NY: Academic Pres*, 376-379.

[14] Gupta, S., Singh, H., Jayanta,Y. (2023). Comprehensive Study on Edge Detection. *Proceedings of the NIELIT's International Conference on Communication, Electronics and Digital Technology* (pp.445-462) 10.1007/978-981-99-1699-3_30.

[15] Madenda, S. (2005). A new perceptually uniform color space with associated color similarity measure for content-based image and video retrieval. *Multimedia Information Retrieval Workshop*, 28th Annual ACM SIGIR Conference.

[16] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700–4708. doi: 10.1109/CVPR.2017.243.

[17] Jiang, J., Wang, C., Liu, X., & Ma, J. (2021). Deep learning-based face super-resolution: A survey. *ACM Computing Surveys (CSUR)*, *55*(1), 1-36.

[18] Arjovsky, M., Chintala, S., & Bottou, L. (2017, July). Wasserstein generative adversarial networks. In *International conference on machine learning* (pp. 214-223). PMLR.

[19] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).

[20] Miyato, T., Kataoka, T., Koyama, M., & Yoshida, Y. (2018). Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*.

[21] Qin M., Mavromatis S., Hu L., Zhang F., Liu R., Sequeira J., and Du Z. (2020) Remote Sensing Single-Image Resolution Improvement Using A Deep Gradient-Aware Network with Image-Specific Enhancement. *Remote Sensing*. 2020; 12(5):758.

[22] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[23] Bottou, L. (1991). Stochastic Gradient Learning in Neural Networks.

[24] Kingma, D. P. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[25] Li, C., & Wand, M. (2016). Precomputed real-time texture synthesis with markovian generative adversarial networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14* (pp. 702-716). Springer International Publishing.

[26] Wang, Y., Bashir, S. M. A., Khan, M., Ullah, Q., Wang, R., Song, Y., ... & Niu, Y. (2022). Remote sensing image super-resolution and object detection: Benchmark and state of the art. *Expert Systems with Applications*, *197*, 116793.

[27] Malakshan, S.R., Saadabadi, M.S., Mostofa, M., Soleymani, S., & Nasrabadi, N.M. (2023). Joint Super-Resolution and Head Pose Estimation for Extreme Low-Resolution Faces. *IEEE Access, 11*, 11238-11253.