Use of AI for Satellite Model Determination from Low Resolution 2D Images

Leon Muratov, Timothy Perkins, Marsha Fox
Spectral Sciences Inc. Burlington, MA
Xuemin Jin
Northeastern University, Boston, MA
Paul LeVan
Air Force Research Laboratory/RDST

ABSTRACT

With the growing number of inactive and active man-made space objects orbiting Earth, space situational awareness (SSA) is becoming an ever more critical element of both national security and the commercial use of space. Accurate tracking and identification of resident space objects (RSO) through space-based or ground-based imagery can provide a means of characterizing potential threats to our orbital assets, and possibly to infer the intent of foreign objects. This work discusses the development of a new satellite identification tool that employs physical model predictions and deep learning neural networks (NN) to increase the quantity of information that can be extracted from these sources of space surveillance imagery.

When observing distant objects in geo-stationary or midcourse orbits, or small satellites in low orbits, unresolved imagery with less than a dozen resolution elements may be the only available optical measurements. Our satellite identification method, the Cognitive Image Recovery Code (CIRC), is designed to evaluate low resolution imagery, including views from multiple locations and temporal light curves generated from unresolved imagery, and to predict the most likely RSO configuration from a list of possible models. For a number of viewing and illumination angles, the tool applies a dense neural network to identify characteristic features of each model and evaluates the most likely match to an unresolved observation. The training imagery for the machine learning algorithm was generated using QUID (QUick Image Display), our fast, first-principles signature simulation code, and a catalog of physically attributed 3D models of various satellites and RSO types. The fidelity of these simulations ensures that the training imagery is both realistic and radiometrically accurate, and the computation speed generates images on-the-fly, allowing an iterative refinement of the model prediction.

1. INTRODUCTION

The growing number of inactive and active man-made resident space objects (RSOs) orbiting Earth requires advances in space situational awareness (SSA) to counteract the increasing potential for collisions, orbital interference and, potentially, active attacks. Identification of RSOs and their payloads provides critical information to agencies responsible for SSA. For an unknown object, passive imaging and temporal variation in radiometry along an orbit provide significant clues to identification. However, for ground-based and even experimental space-based platforms the spatial resolution of imagery is limited due to the range and optical instrumentation limitations. A new capability is needed to expedite modeling of 3D satellite geometries from these low-resolution images.

In this work, we describe the initial development of a new satellite identification tool, Cognitive Image Recovery Code (CIRC). The tool processes low-resolution imagery, including multiple aspects from one or more imaging sensors, and temporal light curves generated from unresolved imagery, and compares it to synthetic imagery generated from a library of 3D satellite models. The objective is to retrieve the most probable satellite configuration and attitude, and to identify anomalous features of the RSO signature. The desired result is a derived 3D model of the satellite and a quantitative assessment of uncertainties with respect to other models in the database. The tool is powered by a deep learning neural net (NN) algorithm to characterize the anticipated variability of RSO signatures under different viewing parameters. The main advantage of CIRC over other satellite identification approaches using AI is that it uses a limited training set for the specific time, geometry, sensors and potential RSOs of the object in question, substantially reducing the time and potential for error of NN training and application to data.

The input to CIRC includes complete imaging scenario information about coordinates, the sensors from which measurements were taken, prior knowledge about likely RSOs, and RSO wireframe models. Fast 3D renderings of wire-frame satellite models allow CIRC to create a scenario specific NN training set of 2D views. The model imaging data is degraded with the proper illumination and sensor resolution for direct comparison with the imagery data. The computed training set is used to train the neural net, by minimizing cross-entropy (metrics of the output quality) of the image set. Once trained, the sensed imagery is processed through the trained NN for classification. During this process, the variances of net weights or conditional probabilities are evaluated to determine whether the model is identified with sufficient confidence. If not, the model can be refined and the entire training process repeated.

In this paper we report our initial findings, that serve as a proof of concept study, and demonstrate that CIRC can select the correct 3D model with a high level of confidence based on a set of low-resolution images. The study assumed that a limited set of viewing configurations and RSO models would be available to the NN. It considered the potential for multiple views of the RSO from two sensor platforms to aid discrimination of 3D features. This initial study ignored atmospheric effects, making the assumption that sensors would be sufficiently atmospherically compensated or located on space-based platforms. An evaluation of the performance of the tool against various levels of sensor noise and image resolutions is presented.

2. METHODS AND PROCEDURES

The training set and test data are produced using the Spectral Sciences Inc. QUID 3D object model [1]. QUID generates radiometrically accurate 2D imagery from 3D wireframes or facet model over any spectral bandpass and optical resolution, and can rapidly generate multiple realizations along an orbital trajectory with evolving attitude. The inputs required for these simulations include the spectral reflectivity of the materials making up the exterior of the object, a geometric shape description, typically a 3D wireframe or facet model, and a model of the environmental illumination. For this effort the only illumination factors considered were direct solar and diffuse Earthshine. The spectral reflectivity of the RSO surfaces is modeled using a modified Beard-Maxwell [2] Bidirectional Reflectance Distribution Function (BRDF) description of each material.

For this initial effort we generated visible band (0.42 to 0.58 μm) training data for a collection of seven 3D object models to represent different types of RSOs. The object categories are generically identified in this paper as 'Cone',' Cygnus', 'Dome', 'Tube', 'Landsat', 'Falconsat', and 'GPS sat'. In addition to satellite models we included four simple object that can be considered as examples of space debris.

The imaging scenario is illustrated in Fig. 1. Each set of training images simulates observations from one of two space-based imaging sensors positioned at different viewing points in space. A set contains 360 orientations for each possible object (rotations around one object axis directed toward the earth center with one-degree increments) as viewed by a sensor. As an example, views at the 200° rotation of all seven objects from the first imaging sensor are presented in Fig. 2. Test images were modeled using the same sensor scenario but a different set of orientations. Some test image projections look quite similar to training, but are visibly distinguishable.
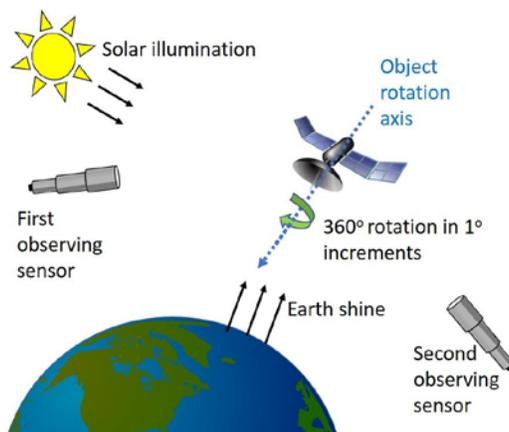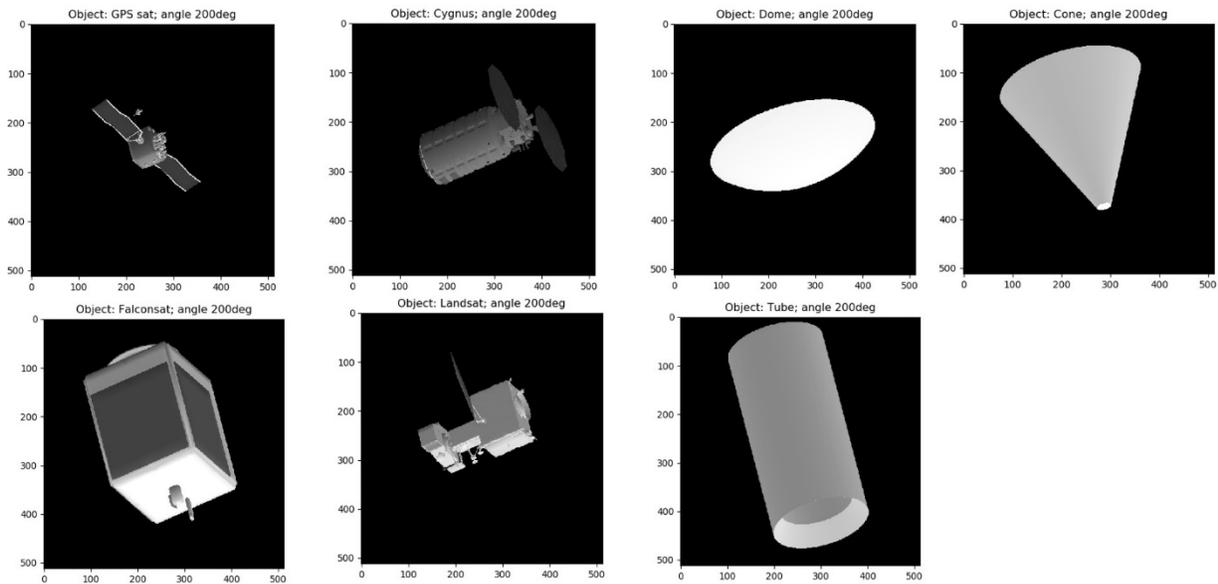

Fig. 1. Simulated observation geometry.

Fig. 2. High resolution images of model objects that were used for the databases generation

To create a NN training set from the high-resolution images, the data was optionally rotated in the image plane, then degraded using a Gaussian blur filter. For this particular test we selected a 10-pixel filter, which corresponds to the diffraction limited resolution of a 20cm telescope at distances from about 100 km to the target. To simulate poor resolution conditions, the images were degraded from 500x500 to 15x15 pixels, effectively reducing the resolvable spatial blur in the images to 3 times the diffraction limited blur diameter. Random noise of varying levels was added. Images from Fig. 2 after processing are shown in Fig. 3. For simplicity, the resulting image intensities were then normalized so that the maximum pixel intensity is uniform across the data set. The noise amplitude was 10% of the maximum intensity. The initial training set includes 25130 images (7 objects, 359 rotations, and 10 different realizations of noise) for each sensor.
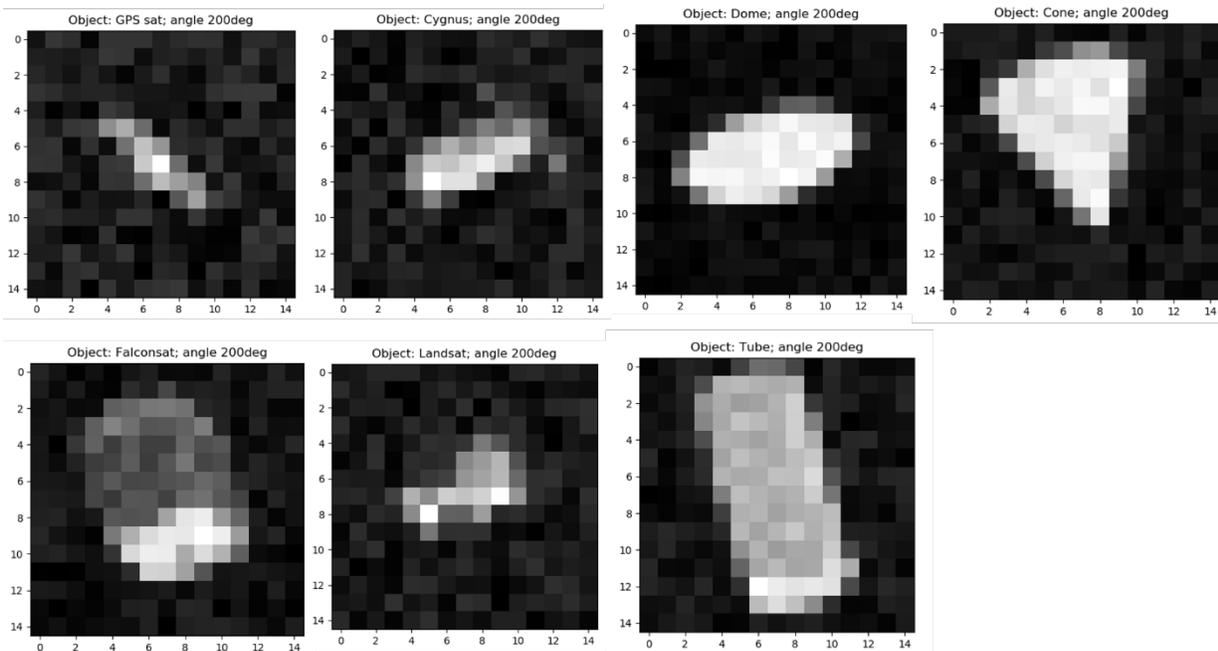

Fig. 3. Degraded images with added random noise. Such images were used for neural network training.

Several sets of test images, were generated independently, by rotating to various angles in the image plane in an anticlockwise direction not applied to the training set. Such additional rotations can be considered as modeled uncertainties in RSO orientation. All test images also have an independent noise realization. Therefore, the testing phase was performed using images that the NN had not encountered during the training cycles. There were 5,026 images in each of the testing sets (2 noise realizations). All sets were randomly shuffled during training and testing phases. It should be noted that the rotations were performed using original high-resolution images to reduce image processing artifacts.

In order to incorporate the two sensor views as input to the NN, the images from each view were combined to form a 'dual perspective' view. Each pair of images were combined in a single frame by concatenating corresponding arrays of pixels. Therefore, each image doubles the number of inputs for the NN. Examples of the images from this set are shown in Fig. 4.
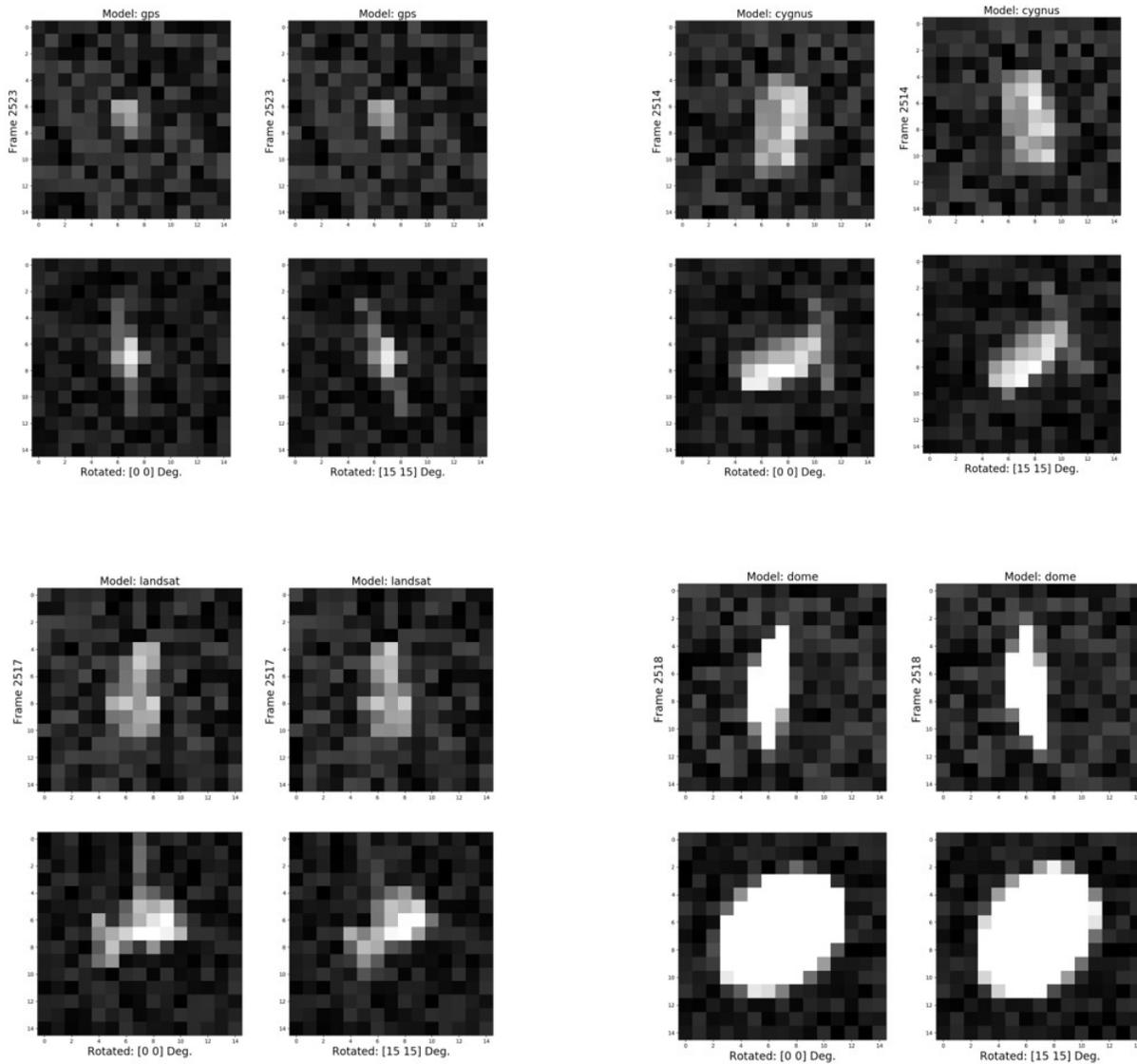

Fig. 4. Four pairs of images of selected RSO Models, with model identifiers listed at the top of each pair. Upper and lower images correspond to observations from the first and second sensor positions, respectively. Each image is also shown before and after applying a 15° rotation.

The final training and testing data sets used in this study included image pairs from two sensor views, with random rotations in the image plane up to a maximum angle of 25°. Random noise and rotations were added to both training and testing sets. For each frame there were 4 such random rotations. The total number of frames in the training set was 100,520 frames (7 models, 359 rotations around main axis, 10 noise realizations, and 4 rotations in the plane). An example of the images is presented in Fig. 5.
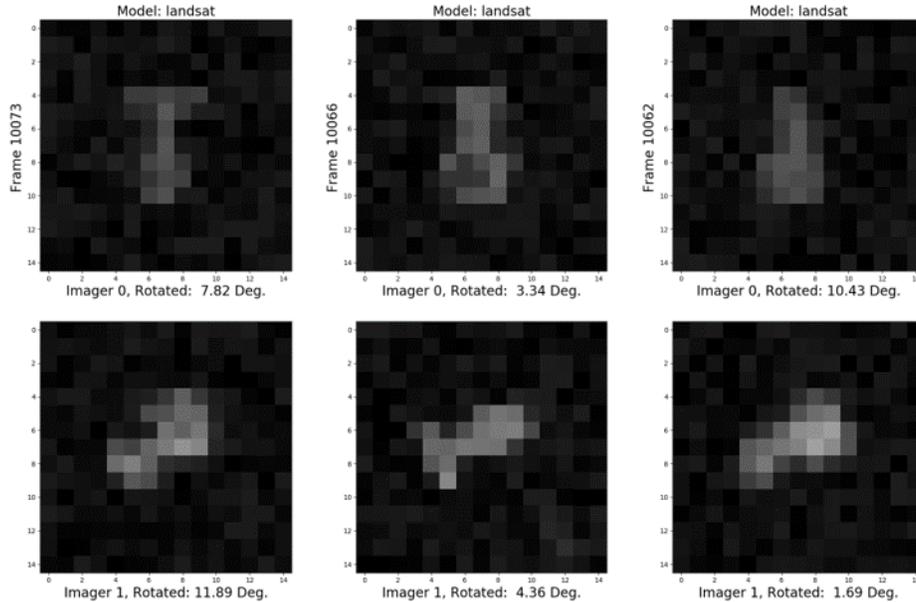

Fig. 5. Randomly rotated in-plane images, each image in pairs rotated independently

For a single sensor, there are 225 inputs for 15x15 pixel images, represented as a single one-dimensional array. For dual-sensor configuration there are 450 inputs, also represented as a single one-dimensional array. With this relatively low number of inputs a simple fully connected NN is sufficient. For this study we used NN with three hidden network layers with 300, 300, and 100 nodes. The neural network was coded using TensorFlow [3] and Python computer language. For minimization, we used the TensorFlow procedure RMSPropOptimizer, which is an implementation of the RMSProp algorithm. This algorithm is a variation of a gradient descent algorithm [4] [5]. We also implemented a mini-batch [Cotter,2011] gradient descent. This algorithm splits the training dataset into small batches and updates model coefficients after each batch is processed. After a few updates a model error is calculated using testing sets. In the results presented in this paper this calculation defines a single iteration. There are also several passages (epochs) over the entire training dataset.

3. TRAINING AND CLASSIFICATION RESULTS

The error rate for the NN training described above is presented in Fig. 6. In each run the training data were split in batches with 500 images in each. Thus, there were 51 batches of data (totaling 25,130 images). Iterations included gradient descent NN optimization using a single batch of data. After every 10 updates, we used test data to determine the error rate and add a point to the graph shown in Fig. 6. The entire process was repeated 15 times. Thus, there are 77 iterations in each run. Multiple lines in Fig. 6 show the results of several training runs. The variability of learning rates is a result of randomness in the initial values of the NN weights. The NN error rate quickly drops to zero after about 15 iterations, meaning that the selected level of noise, resolution, and variety of models form a problem that is easy for the algorithm to solve.
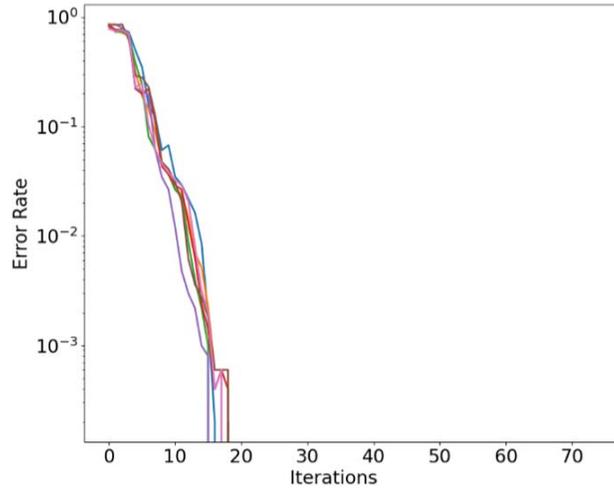
Fig. 6. Results of the NN testing for one imager with noise variations in training and testing sets. One iteration includes ten cycles of optimization with mini-batches of data and one error rate test.

In Fig. 7, we compare error rate of results for single and dual perspective imagers configurations for a 5° rotation of the testing images in the image plane.
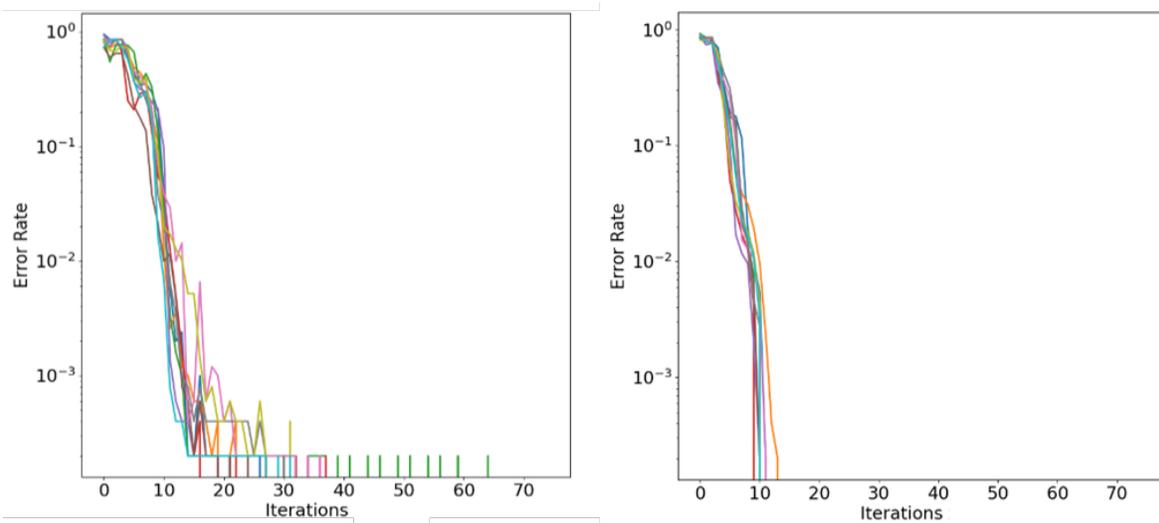

Fig. 7. NN error rates for 5° rotation, the left panel a single imager, the right panel two imagers. The NN is trained using unrotated images and tested on images rotated by 5°.

For a single imager the error rate converges a bit more slowly than in Fig. 6, but it still gets to the zero-error rate. Dual perspective imaging results in more rapid convergence to a zero-error rate. It is interesting to compare the cost functions minimization during the training of the two cases, shown in Fig. 8. The minimization process is independent of testing, and is conducted on unrotated images only. For the dual perspective, the cost function exhibits a more regular behavior and drops almost an order of magnitude lower. Oscillations of the cost function for the single imager can be removed by additional regularization terms in the minimization process.
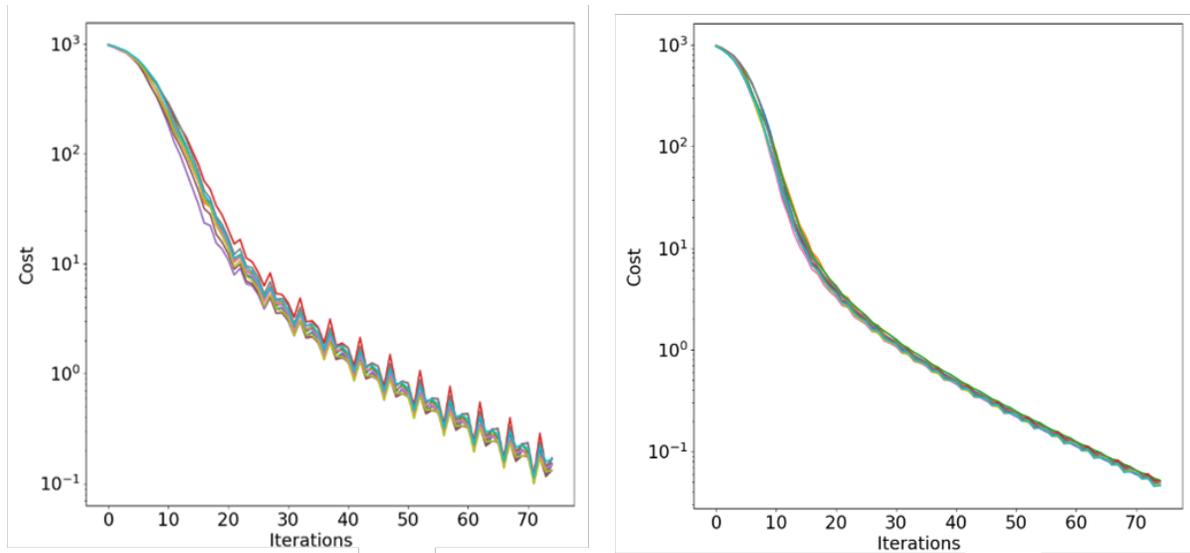
Fig. 8. Decays of cost functions in training processes, the left panel for a single imager, the right panel for two imagers.

Similar testing using 10º and 20º rotations of test objects from the training set orientation also demonstrated the advantage of dual projections. However, the benefits are significantly diminished, especially for 20º rotations. Fig. 9 demonstrates that the error rate usually stabilizes around 8% for 20º rotations. An interesting feature of the plots in Fig. 9 is a drop in the error rate at around 10-15 iterations, followed by an increase. This is a symptom of NN overfitting. It occurs when the machine learning algorithm captures noise patterns of a training data set. When features in the test set are quite different (rotated in our case), a noise pattern becomes more important for the algorithm.
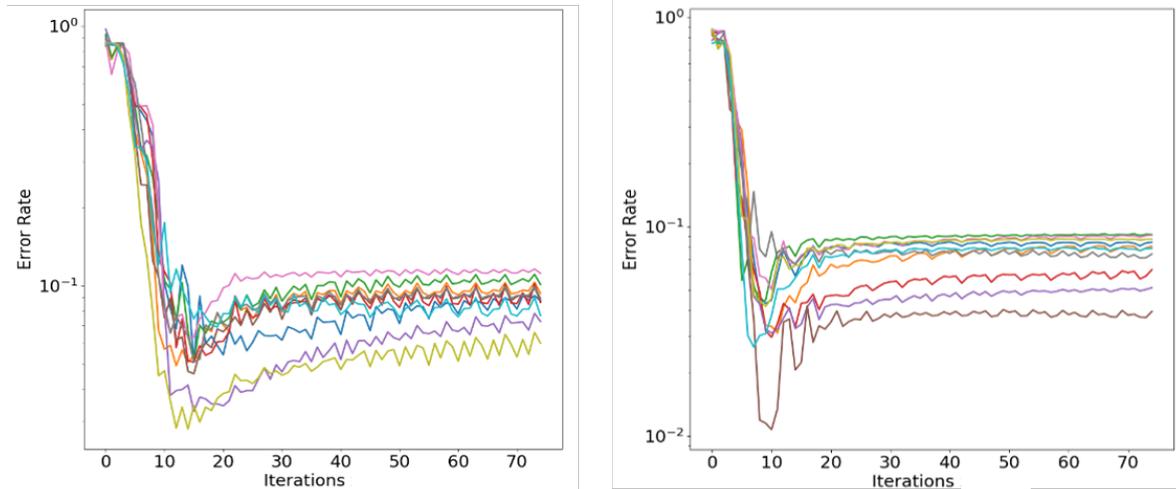

Fig. 9. NN error rate for a single (left) and dual-perspective observation (right). The NN was trained using unrotated images and tested on images rotated by 20°.

The results presented in Fig. 10. are for random rotations with maximum angle of 25º. As expected, addition of random rotations of both the training and the testing set improves the NN performance. The single imager performance probably can be improved by additional regularization, but the dual imager performed exceptionally well and yields a perfect classification rate after fewer than 40 training cycles.
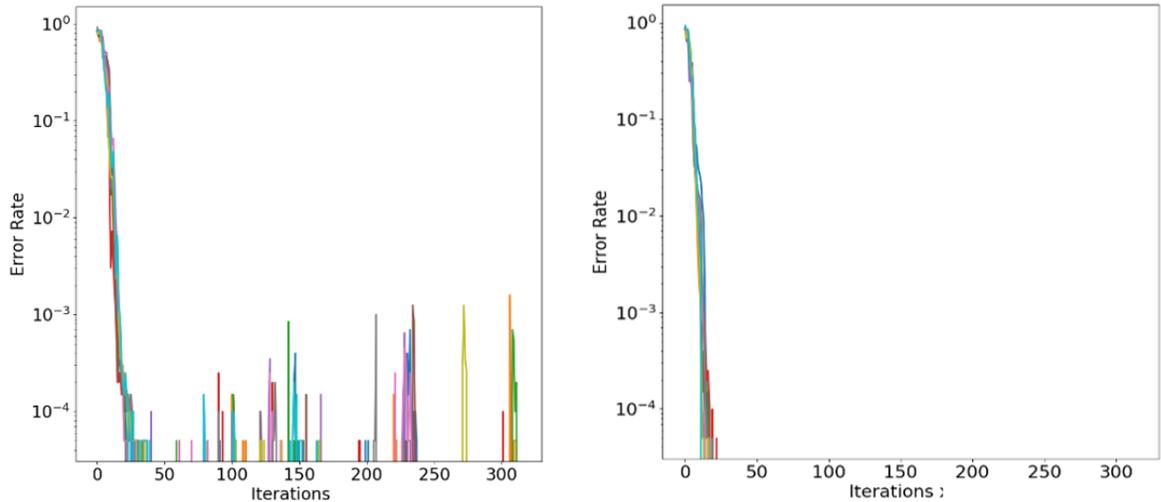
Fig. 10. NN error rate when trained with random in-plane rotations, the left panel for a single imager, the right panel for two imagers. With four random in-plane rotations the training totaled 100,520 images.

4. SUMMARY AND FUTURE DEVELOPMENT

To summarize, this study demonstrated that NN algorithms can successfully learn to identify RSOs from noisy, low resolution images. The results quantify the classification accuracy for imaging scenarios using single and dual-perspective observations. The results can be generalized to multiple views, however increasing number of input pixels will eventually call for more complex analysis methods, such as the use of convolutional neural networks to extract key features. Future development work will apply this technique to a larger library of RSO models, and will investigate methods to modify and adapt RSO models within the NN training cycle. Methods to estimate the confidence of the RSO classification will also be investigated.

5. ACKNOWLEDGEMENT

6. REFERENCES

[1] T Perkins, R Sundberg, J Cordell, Z Tun, and M Owen. Real-time Target Motion Animation for Missile Warning System Testing, *Proc. SPIE* Vol 6208, Kissimmee, Florida, April 2006.
[2] WC Shemano, W F Lynn, D Shemano, and TA Davis. Modification of the Maxwell-Beard bidirectional reflectance distribution function, *Proc. SPIE* 3784, 240–248 1999.
[3] TensorFlow, https://www.tensorflow.org/ 2019.
[4] S. Ruder. *An Overview of Gradient Descent Optimization Algorithms*, arXiv preprint arXiv:1609.04747, 2016.
[5] A. Cotter, O. Shamir, N. Srebro, K. Sridharan. *Better Mini-Batch Algorithms via Accelerated Gradient Methods*, in Advances in Neural Information Processing Systems 24 (NIPS): 1647-1655, 2011.