Object Characteristic Determination Using Brightness Measurements

Pace Balster, Gabrielle Jones, Gavin Hofer, Daniel Newsom
Katalyst Space Technologies, LLC

Carolin Frueh
School of Aeronautics and Astronautics, Purdue University

ABSTRACT

Determination of space object characteristics is an increasingly important capability as the space domain becomes more congested and complex. The amount of collected Space Situational Awareness (SSA) data has grown considerably in the past years, straining existing analysis and processing techniques. While an uprecedented amount of data is being collected, the quantity of information gathered on individual objects can differ significantly. This variation necessitates the use of automated and scalable characterization techniques that can adapt to inconsistencies in the density or sparsity of observation data. This technical paper constitutes an advancement in the ability to predict space object type by utilizing a model that is extensible to different numbers of non-resolved observations, even when these observations are relatively few and are collected over short time durations. An investigation into an automated machine learning technique for object type prediction is conducted. Aggregate statistical features from optical observations as inputs to the model are contrasted with features generated from a partial light curve inversion approach. The methods are compared independently and in conjunction with each other, with the combined aggregate and partial inversion feature set shown to offer improvement over either in isolation. Techniques like the ones presented in this paper would enhance the toolset available to the SSA community for object characterization, thereby improving understanding of space objects to better inform decision-making in an increasingly dynamic space environment.

DISTRIBUTION A: Approved for public release; distribution is unlimited.
Public Affairs release approval #AFRL-2023-4203

1. INTRODUCTION

A rapidly evolving space domain spurred by a dramatic surge in launch rates, the deployment of large satellite constellations, and more complex space operations has underpinned a need for timely understanding of space object characteristics. While global sensor networks are collecting more Space Situational Awareness (SSA) data than ever, challenges still persist in the ability to fully exploit this information. Human-made space objects exhibit large distances to Earth-based observers, ranging from several hundred to several thousand kilometers. The objects themselves are small relative to those scales, being on the order of meters. As a result, space objects, outside the lowest Low Earth Orbit (LEO) region are not able to be observed in a resolved fashion. Additionally, while some high-priority space objects are tracked near-continuously, the focus often comes at the expense of adequately characterizing other objects. Consequently, when aiming to collect information on a large number of objects routinely, the data available is often non-resolved and sometimes sparse, reflecting the trade-offs made in prioritization. This necessitates the development of methods that can adequately characterize space objects, even when faced with few non-resolved observations.

Non-resolved optical observations are composed of astrometric and photometric information. Astrometric information related to object angular position and rates can be used straightforwardly for orbit determination; however, for further characterization, an object's reflected sunlight over time offers additional information. In the case of passive observations, a light curve refers to the reflected brightness of an object when arranged in a time series. The brightness is dependent on many factors including object orientation, shape, composition, and solar phase angle. Though analyzing a light curve is an inherently multivariate problem, the encoded information within it offers an opportunity to derive various characteristics from the reflecting body such as size, shape, attitude, mass, and materials. Light curves have been explored in astronomy, and significant research has been done in expanding light curve inversion (LCI), a technique for determining shape and attitude from the light curve, to human-made objects.

In general, one can discern full shape inversion techniques as a variation on geometry-based methods derived from asteroid light curves [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. These methods have been shown to be effective on convex objects. Unfortunately, most human-made artificial objects show sharp edges, and significant concavities, subject to self-shadowing effects. Alternatively, attitude-only inversions have been developed [11, 12, 13, 14, 15, 16]. Attitude inversion often suffers from simplifying assumptions, such as a cylindrical shape or stable, single-axis, or mainly single-axis rotation. Only a few combined shape and attitude approaches have been explored [17, 18].

Because of these shortcomings, shape-surface determination has been approached using filtering approaches [18, 19, 20, 21]. In order to provide the necessary initial conditions and to limit the vast parameter space to allow for convergence, the problem has been made tractable by using simple a priori shapes such as cubes, and cylinders and obtaining a weighting in a mixture of experts, filter banks or multiple hypotheses techniques. The addition of orbital information in order to narrow down ambiguities via the use of orbital perturbations has not yet shown significant improvements.

Due to the challenges discussed, light curve inversion methods often require the aforementioned limiting assumptions that reduce their generality to a subset of objects, such as convex objects with uniform material, not adequately addressing the need for better characterization of human-made objects across different sizes, shapes, and orbital regimes. While the full inversion problem for shape and attitude is extremely difficult, methods for partial inversion, sometimes referred to as "Blob Identification" aim at partial inversion results. With Blob Identification (Blob ID), not a full inversion is sought, but rather the extraction of certain features of the object, often in a trade-space analysis against multiple plausible options, similar to a mixture of expert approaches. This greatly diminishes the assumptions that need to be made, while providing a focus on the more readily observable characterization features.

Furthermore, Machine learning (ML) methods have shown applicability across a variety of domains, including the characterization of space objects [22, 23, 24]. Machine learning can be subdivided into classical and deep learning methods. Classical methods, as defined here, are a set of algorithms and statistical models that enable computers to learn relationships between features in a data set to make predictions on new data. Specific classical methods include Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines. These approaches are particularly advantageous when the training data set is relatively small and lacks complexity. New advancements in machine learning, including deep learning, have shown promising results across a large number of domains including computer

vision [25] and natural language processing [26]. Deep learning methods typically refer to multi-layered artificial neural networks that comprise hidden layers of interconnected nodes to expose complex feature relationships by learning on a large training data set [27]. In general, both classical and deep learning ML methods are advantageous due to their ability to handle non-linearities, high dimensional data, and their tolerance to noise. Additionally, ML models can be highly efficient to run after they have been trained. These traits present applicability for the multi-variate object characterization problem discussed earlier.

ML approaches have been previously investigated to assist in extracting shape and attitude information from the inversion of light curves. Agathanggelou et al. [22] found that the best performing classical model, Random Forest, performed just as well as a deep learning approach, specifically a Long Short Term Memory (LSTM) model for predicting the operational status of resident space objects (RSOs) in LEO. The term "track" was defined as all the time-series measurements for a single observed pass overhead of a given satellite. For a single track, the Random Forest model achieved an accuracy of $83 \pm 1\%$ for LEO and $90 \pm 1\%$ for GEO.

Other efforts have been made to apply supervised and unsupervised machine learning techniques to partially solve the light curve inversion problem. Furfaro et al. [28] used deep learning algorithms to learn the inverse relationship between light curves and shape. Using a supervised approach, they were able to classify different shapes of rocket bodies using a Convolutional Neural Network (CNN). Methods for unsupervised learning like clustering and dimensional reduction were investigated for satellite classification but they were not conclusive on their own [22, 28].

While there already exists published work on a variety of methods for RSO characterization using non-resolved optical observations, these works are often confined to either pure machine learning approaches or direct methods such as light curve inversion (LCI) techniques to extract shape or attitude and are typically evaluated on long time-series observation and data.

This paper explores a supervised machine learning method for predicting the type of observed object using input features derived from astrometric and photometric data in an electro-optical (EO) data set composed primarily of GEO objects. A study of feature engineering methods is conducted to understand the predictive power of partial inversion (Blob ID) as an input to the model in comparison to purely aggregate statistical quantities. The same Random Forest architecture is employed across all methods to isolate the effects of changing the input features. Object type categories are delineated into four categories: active payload, inactive payload, rocket body, and debris. The methods are evaluated on object EO tracks with observations ranging from 10-15 per track. The ability to function without the need for many observations enhances extensibility, allowing the methods to be applied to various observation sites and objects that are only briefly observed. While the Blob ID methods presented here perform better on longer tracks due to their ability to fit periodicity in the data, this paper shows there may still be predictive power of the partial inversion outputs in addition to the aggregate statistical features for short tracks.

2. METHODS

The direct observables from non-resolved electro-optical observations of RSOs are astrometric quantities (i.e., right ascension and declination), along with the brightness value. These observables are calibrated using the background stars present in the same frame. For astrometric reduction, the Gaia catalog is adjusted with the Tycho proper motions. The Gaia catalog is used directly for brightness calibration. During calibration, stars spatially close to the observed object are assigned more weight to accurately scale the stars within the same field of view.

Brightness, expressed as calibrated apparent magnitude, depends on several factors. These include the satellite's shape, albedo, material composition, attitude, the angle between the observation vector and the sun vector (solar phase angle), and the distance from the observer and object to the sun. Given that brightness fluctuates based on these factors, the change in brightness over time, or the light curve, can be used to further characterize the object. When the light curve is used in conjunction with the astrometric quantities, an even more comprehensive characterization of the object can be attained.

This study explores the effectiveness of predicting object type based solely on electro-optical observation data. The primary data source used for training, testing, and evaluation of the methods is a Wide Area Search (WAS) EO data set provided by the 15^{th} Space Surveillance Squadron (15^{th} SPSS). The 15^{th} SPSS operates the Maui Space Surveillance Complex (MSSC), located at the summit of Haleakala on the island of Maui, Hawaii, host to multiple telescopes.

2.1 Data Features

In this study, "tracks", sequences of electro-optical observations on a given object, are analyzed to predict the object's class (active payload, inactive payload, debris, or rocket body). Because both astrometric and photometric data are used for input features to the model, a track, as defined here, is considered to be made up of observations containing both astrometric and photometric information. Tracks varying from 10 to 15 observations are used, with durations on the order of several seconds. All of the tracks analyzed are produced by objects ranging from Medium Earth Orbit (MEO) to Geosynchronous Orbit (GEO), with the majority of tracks being from GEO objects. The following values from each observation are used for analysis: Time in seconds from the first observation in the track; brightness, expressed as calibrated apparent magnitude; right ascension and declination in degrees; right ascension and declination rate-of-change in degrees per second.

2.2 Labeling the Data

Celestrack is an open-source website that maintains a satellite catalog (SATCAT) [29] containing information about RSO orbital details, owner/operator information, object tags, and current operational status. From the SATCAT, object tags (payload active, payload inactive, debirs, and rocket body) are collected for all NORAD IDs in the catalog. The NORAD IDs with the corresponding tags are mapped to the 15^{th} SPSS data set to enable evaluation of the supervised ML classification approaches.

SATCAT Operational Status

Operational Status	Descriptions
+	Operational
-	Nonoperational
P	Partially Operational *Partially fulfilling primary mission or secondary mission(s)*
B	Backup/Standby *Previously operational satellite put into reserve status*
S	Spare *New satellite awaiting full activation*
X	Extended Mission
D	Decayed
?	Unknown

*Active is any satellite with an operational status of +, P, B, S, or X.
Active status does not require power or communications (e.g., geodetic satellites)

Fig. 1: The Celestrack SATCAT operational status used to define active and inactive payload labels [29]

An important consideration is the accuracy of the labels. When comparing the accuracy of two independent sources of labels (Seradata and Celestrak) the overall agreement was 95% on average between the label sources [22]. The accuracy of the model can only be as accurate as the labels themselves, so this provides an upper bound for model performance.

2.3 Classifier Training and Validation

Verification and validation are essential for quantifying the robustness and generalization of ML models to new cases and data. To test this, it is typical to maintain a hold-out set of data that the model has not used or seen before to test the performance of the model when introduced to new cases. In this study, a method called Repeated Stratified K-fold Cross-Validation [30, 31] is used to evaluate each classification approach. Repeated Stratified K-fold Cross-Validation is a technique that provides a robust measure of model performance by partitioning the data into 'k' randomized subsets or 'folds' multiple times while ensuring the class distribution of each fold mirrors the overall distribution. For each testing cycle, the classifier is trained and validated 'k' times, each time using a different fold as the validation set (also known as a "holdout" set), and the remaining folds as the training set [30].

To increase the reliability of the performance estimates for the model, multiple testing cycles are carried out with randomized fold selections in each cycle. For each of these testing cycles a new randomized subset with evenly distributed class labels is selected prior to splitting into folds. Additionally, each fold is stratified to inherit the balanced class label distribution. Balanced subsets are used both to address difficulties in training predictive models on unbalanced datasets, and to improve the interpretability of the performance results. For performance evaluation, the mean of each of the performance metrics is calculated across all folds and testing cycles. All performance results in this study are produced using 5 folds, repeated across 100 test cycles, for a total of 500 individual classifier fits.

Notably, the folds are produced by selecting tracks by catalog ID, so that each individual RSO is isolated to one fold for cross-validation. This approach is done to prevent a form of data leakage that could occur if the model is evaluated using tracks produced by any of the same objects found in the training set.

The Random Forest model was selected as the model of choice for this study due to its comparable performance to deep learning methods for object type prediction. Additionally, selecting a relatively predictable model, such as Random Forest, allowed for the isolation of the predictive power of different input features including Blob Identification from that of the aggregate statistical features.

2.4 Classification Methods

Three different methods are compared: Method 1, the *Aggregate Random Forest Classifier* employs aggregate statistical quantities computed from the tracks as features for a Random Forest classifier. Method 2, the *Blob Identification* method, involves a direct optimization approach that fits several idealized mathematical models of different types of objects to the actual light curves using gradient descent. The normalized mean squared errors (MSEs) of these models are then used along with boolean flags for whether the model converged as features for the Random Forest classifier to predict the object type. Method 3, the *BlobID-Aggregate Random Forest Classifier*, is a combined approach. It uses the aggregate track quantities along with the Blob Identification residuals and boolean convergence flags as features for a random forest classifier.

For all three methods, the RandomForestClassifier from scikit-learn [30] is used for implementation. The Random Forest classifier [32], a machine learning algorithm, uses multiple decision trees trained on bootstrapped subsets of the training data. The final prediction is the majority vote of class probabilities, averaged over all tree predictions. This method reduces overfitting common in decision tree models while maintaining their strong predictive power.

2.4.1 Method 1: Aggregate Random Forest Classifier

To transform time-series track data into a set of interpretable input features for the Random Forest classifier, the following statistical aggregate quantities are computed for all six channels over each track: mean, maximum, minimum, and variance of the raw values as well as the back-differences of the raw values. All together, there are 48 aggregate features available as inputs to the Random Forest. The architecture of the Aggregate Random Forest Classifier is shown in Fig. 2.

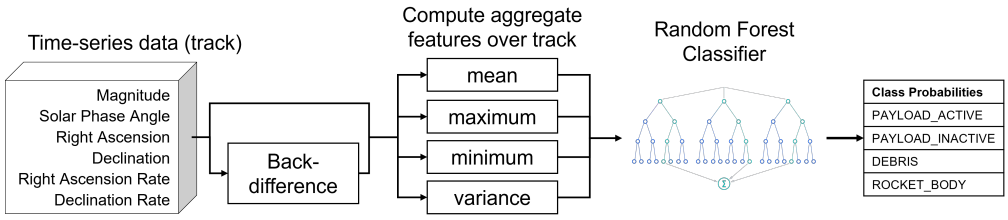

Fig. 2: Aggregate Random Forest Classifier

These aggregate features are used instead of the raw values from the observations such that tracks of varying lengths and time intervals can be used, and to reduce sensitivity to random fluctuations. In addition to these statistical features, features extracted from frequency domain analysis were also investigated. Ultimately, these features are not used due to difficulties in extracting periodic behavior from shorter tracks.

2.4.2 Method 2: Blob Identification

In lieu of a full inversion, the equations governing the reflected light can also be used to ascertain more generalized, incomplete characterization information, supporting the association of objects in different categories. Such information, often dubbed Blob Identification (Blob ID), can stand on its own or be used, as in the case of this paper, in conjunction with machine learning methods.

The Blob Identification method relies on the principle of finding relations that fit a certain set of generalized albedo-shape and frequency relations that are in agreement with the proposed satellite categories. Based on the fit, which often is poor, weights can be calculated, similar to a mixture of expert fashion [20]. The difference is that no full-scale models are fit, but sub-facets in proportionality relations rather than absolute models. Furthermore, no full attitude profiles are matched, but only frequencies are extracted.

The principal relations that are used are the reflection properties of a flat plate and round surface reflection. The irradiance I_{plate} from a flat plate defined by its surface normal is captured by the following proportionality relation:

$$I \propto \cos\theta_{\text{S}}(p \cdot \cos\theta_{\text{O}} + q \cdot \delta(\theta_{\text{S}} - \theta_{\text{O}})\delta(\phi_{\text{S}} - \phi_{\text{O}} + \pi)), \tag{1}$$

where θ_{O} is the out-of-plane angle between the surface normal and the observer direction, and θ_{S} to the sun direction, respectively. ϕ_S and ϕ_O are the corresponding in-plane angles, guaranteeing the glint condition. In reality, the sun is an extended source and there is a 0.5-degree deviancy in the glint direction allowed, however, solar panels might also be slightly misaligned making this difference negligible. p and q are the albedo-area relations of Lambertian to specular reflection properties. For Blob ID, the relation between p and q is not fixed; also the condition that they have to add up to 1 is not enforced, hence acting as a general scaling parameter accounting for the entirely coupled area-albedo relation.

With the flat-plate proportionality relation, any box-bus configuration and solar-panel (single or dual wing) relation can be modeled for a fit to an actual light curve. For a controlled box-wing satellite, the solar panels are oriented towards the sun, so one flat plate orientation has a necessary orientation of angle between the normal and the sun $\theta_{\text{S}} = 0$ and, hence, the angle between the normal and the observers is the same as the overall phase angle $\theta_{\text{O}} = \alpha$. For the reflectivity, it is irrelevant if it is a two or one solar panel configuration. The eventual blockage of the view to a panel to the observer by the bus is equivalent to a different albedo-area and hence part of the scaling in the fit. Maximally three surfaces of a box-shaped bus are visible, constrained in a relative 90-degree configuration. Only two sides at most are illuminated for a controlled satellite with an approximately 90-degree angle to the actuating panels, which are sun-oriented. Side 1, which is just offset by the solar-panel rotation angle β to follow the sun, leads to a $\theta_{\text{S}} = \beta$ and $\text{O} = \alpha - \beta$.

A satellite can go through several instances of rotational and tumbling motions when no longer attitude stabilized [13]. Fewer restrictions can be applied, as the orientation towards the sun is lost for any surface. A constraint that can be enforced at any time is the one of the overall phase angle $\alpha = \theta_{\text{S}} + \theta_{\text{O}}$. This constraint can be enforced for a cube configuration representing the box, where at most three sides are visible at any time. The solar panels are assumed to be at a random actuation angle and in a 90-degree configuration relative to the box. In general, enforcing the relative relations via the phase angle has not shown to provide significant improvements compared to just keeping the box-wing configuration without the phase angle constraint.

The stabilized and non-stabilized box-wing satellites are most distinct in their relative synodic variation that is observable in the light curve, corresponding to the different attitude profiles. For the analysis, the periodogram method is far superior in terms of robustness in the presence of noise over a Fourier analysis. Box-wing satellites in terms of their relative inertia distribution and relative rigidity of the object itself, tend to result in synodic periodic patterns, which repeat over short time periods [33]. As such, an additional constraint is applied in terms of detecting distinct periods over-weighing a non-stabilized hypothesis relative to a stable box-wing one.

A large group of objects are cylindrical in shape. This includes spin-stabilized objects and their debris following end-of-mission and the larger group of upper stages, some of which also experience a spin-up during deployment. Spin-stabilized satellites following end-of-life will continue to maintain spin through momentum conservation. Although eddy currents contribute to a slowing of the spin, collected torques create rotations around other axes. Recently passivated spin-stabilized objects are indistinguishable from still spin-stabilized ones over short observation times in the light curves.

Cylindrical objects are characterized via the reflection off the mantle, which is orthogonal to the spin axis identical to the spherical reflection, leading to the following irradiance relation:

$$I_{\text{sphere}} \propto \cos\theta'_{\text{O}} p(sin(\alpha) + (\pi - \alpha)\cos(\alpha) \tag{2}$$

where θ'_{O} is the angle of the cylindrical axis alignment relative to the observer direction, and α is the phase angle. The top and bottom face, which are flat plates, are governed by Eq.1. Nozzles are neglected, as they are absorbed into the mantle albedo-area-relations.

For a cylindrical spin-stabilized object, the spin axis aligns orthogonal to the sun, bringing Eq.2 to its full bearing. The angle to the observer θ'_O is then fully determined via the orbital plane and can be computed directly from the state of the object at each point in time during the light curve. A cylindrical uncontrolled object does not maintain this relation. Therefore the combined proportionality relation of Eqs.2 and 1 are combined with an open relation of the angle θ'_O, only enforcing the orthogonality of the mantle relative to its flat top and bottom.

For a cylindrical uncontrolled object, significant attitude motion does take place, which clearly discriminates it from a spin-controlled object of the same shape. The light curves of spin-stabilized cylindrical objects are extremely flat, due to the gentle cylinder mantle, which can be detected using the periodogram analysis. Freely tumbling cylinders fit the curved surface relation, Eq.2, while also showing significant periods even over short time spans. It has to be noted that there generally exists non-uniform distribution of mass, e.g., upper stages with slosh and objects with a heavy motor on one end and a (mostly) empty hull at the other. As such, the attitude motion is often more complex and has more than one dominant period present. Again, the periodogram method can be used to over- or underweight hypotheses that stem from the intensity proportionality fits.

To predict object classes using the Blob Identification method, first, the irradiance of the observation is computed :

$$I = I_0(10)^{-0.4m}, \tag{3}$$

where I_0 is the reference irradiance, and m is the calibrated apparent magnitude. Note that since absolute scaling is not required for Blob Identification, I_0 can be set to 1.0 while preserving the relative change of irradiance over time.

Four different models based on the relationships discussed above are fitted to the light curve to minimize the mean squared error (MSE) of irradiance. The four models represent stabilized and non-stabilized versions of both a box-wing satellite and a cylinder. The final MSEs of a batch-optimized fit are then used along with boolean flags for whether or not each model converged as features in a random forest classifier to predict the object class. The Random Forest classifier is used to learn the correlation of these features to object class from the data and get an estimate of the full predictive power of these residuals. Figure 3 shows how the results from Blob ID are used as input features for Random Forest to predict the object classes.

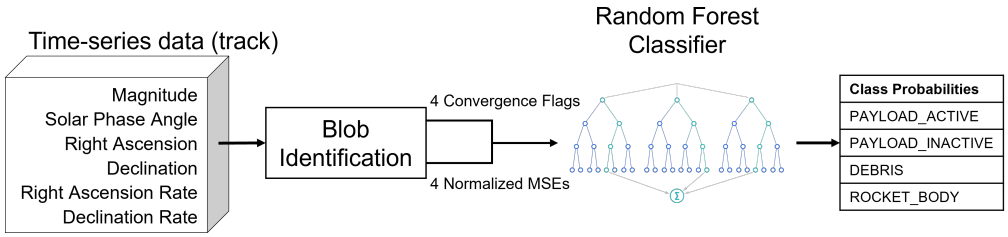

Fig. 3: BlobID Random Forest Classifier

2.4.3 Method 3: BlobID-Aggregate Random Forest Classifier

The final classification method, the BlobID-Aggregate Random Forest Classifier, is a combination of the previous two methods. This method uses all of the aggregate quantities computed for each track as discussed in 2.4.1, in addition to the normalized mean squared errors and boolean convergence flags of the Blob Identification models, as features for a random forest classifier. The purpose of this approach is to explore the additional predictive power of the Blob ID residuals beyond that of the aggregate statistical quantities alone.

With the 48 statistical aggregate quantities, 4 MSEs, and 4 convergence flags, the Random Forest in this approach takes in a total of 56 input features. The architecture of the BlobID-Aggregate Random Forest classifier is shown in Fig. 4.

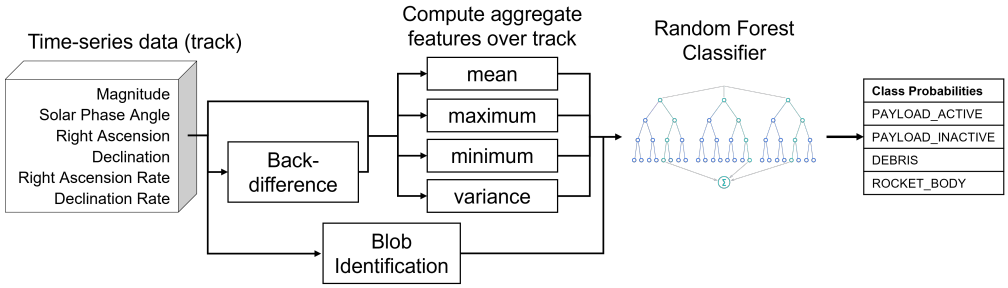

Fig. 4: BlobID-Aggregate Random Forest Classifier

2.5 Performance Metrics

The three classification methods described in section 2.4 are evaluated using the performance metrics in Table 1. All of these metrics are used to evaluate both the performance in distinguishing individual classes from all the others (one vs all), as well as combined performance. Macro averaging, a method that finds the overall metric by averaging the metric over all class types without considering class imbalances [30], is used to determine the combined precision, recall, and f1 score.

Table 1: Machine Learning Performance Metrics

Metric	**Description**
Precision	The proportion of positive predictions that are true positives.
Recall / True Positive Rate	The proportion of positive cases that are correctly predicted as positive, also known as sensitivity.
F1 Score	The harmonic mean of precision and recall. F1 score provides a balance between the two metrics.
True Negative Rate	The proportion of negative cases that are correctly predicted as negative, also known as specificity.
Balanced Accuracy	The average of the true positive rate and the true negative rate.
Accuracy	The proportion of correctly labeled predictions among all predictions.

3. RESULTS

As described in 2.3, repeated stratified k-fold cross-validation is used to evaluate the performance of each classification method. All results presented in this section are produced using the averages of out-of-sample performance across all folds and testing cycles. Table 2 shows a comparison of the three classification methods explored, in terms of balanced accuracy for individual classes (one versus all), as well as overall accuracy. Balanced accuracy is selected over regular accuracy for individual class predictions because the regular accuracy would be artificially inflated due to the class imbalance in the one versus all classification problem. For the overall accuracy, balanced and regular accuracy are equivalent because a balanced subset is used for evaluation. Note that for the balanced accuracy on individual classes the baseline corresponding to no predictive power is 50%, and for the overall accuracy on four classes the baseline is 25%.

Table 2: Individual Class Balanced Accuracy and Overall Accuracy for each Classification Method

Classifier Performance Comparison (Balanced Accuracy)			
Class	Aggregate	Blob Identification	BlobID-Aggregate
Active Payload	78.1%	58.2%	78.7%
Inactive Payload	79.8%	55.5%	80.1%
Debris	85.2%	53.5%	85.1%
Rocket Body	80.3%	51.9%	80.3%
Overall Accuracy	71.3%	32.2%	71.6%

As shown in Table 2, the Blob Identification classifier performs best when distinguishing active and inactive payloads from other classes, while the classifiers that use statistical aggregate quantities perform better on debris and rocket bodies. Note that the Blob Identification method is most effective on longer tracks, as it relies on phase angle dependencies and periodic temporal dynamics. This study utilized tracks of less than 10 seconds in duration, with phase angles spanning less than 150 arcseconds. Therefore, it is noteworthy that the Blob Identification method still demonstrates some predictive power when applied to these tracks. When combined with the statistical aggregate features, Blob Identification only provides a marginal performance increase, suggesting that the useful information provided by Blob Identification for predicting object class mostly overlaps with that provided by the statistical aggregate features. To ensure that the increase in performance observed with the addition of Blob Identification is not due to random variation, the repeated k-fold cross validation process is repeated 10 times. The standard deviation over these 10 runs in macro-averaged balanced accuracy is 5.45×10^{-3} %. The standard deviation of the overall accuracy over these 10 runs is 1.59×10^{-2} %. More detailed reports of each classifier's performance are shown in 3.1.

3.1 Classification Reports

Tables 3, 4, and 5 show the performance of the three classification methods outlined in 2.4. The selected performance metrics are precision, recall, f1 score, and balanced accuracy, as well as overall accuracy for multi-class performance. For precision, recall, f1 score, and overall accuracy, the baseline (corresponding to no predictive power) is 0.25. For balanced accuracy, the baseline is 0.5.

Table 3: Aggregate Random Forest classifier performance

Aggregate Classifier Performance				
Class	Precision	Recall	F1 Score	Balanced Accuracy
Active Payload	0.845	0.599	0.701	0.781
Inactive Payload	0.631	0.741	0.681	0.798
Debris	0.716	0.811	0.761	0.852
Rocket Body	0.712	0.701	0.706	0.803
Macro Average	0.726	0.713	0.712	0.809
Overall Accuracy	0.713			

Table 4: BlobID-Only Random Forest classifier performance

BlobID-Only Classifier Performance				
Class	Precision	Recall	F1 Score	Balanced Accuracy
Active Payload	0.357	0.407	0.381	0.582
Inactive Payload	0.345	0.300	0.321	0.555
Debris	0.301	0.316	0.308	0.535
Rocket Body	0.280	0.263	0.271	0.519
Macro Average	0.321	0.322	0.320	0.548
Overall Accuracy	0.322			

Table 5: BlobID-Aggregate Random Forest classifier performance

BlobID-Aggregate Classifier Performance				
Class	Precision	Recall	F1 Score	Balanced Accuracy
Active Payload	0.850	0.611	0.711	0.787
Inactive Payload	0.639	0.742	0.687	0.801
Debris	0.715	0.810	0.760	0.851
Rocket Body	0.711	0.702	0.707	0.803
Macro Average	0.729	0.716	0.716	0.811
Overall Accuracy	0.716			

As seen in Tables 3, 4, and 5, both classification methods that incorporate the statistical aggregate quantities significantly outperform the method that only uses the outputs from Blob Identification as input features. Including the outputs from Blob Identification results in a marginal increase across all measured performance metrics indicating some unique predictive power in the method.

3.2 Precision-Recall and Receiver Operating Characteristic Curves

Precision-recall (PR) and Reciever Opterating Characteristic (ROC) curves are both generated by defining a set of varying classification thresholds from 0 to 1. A prediction is flagged as a specific class if the classifier output probability for that class exceeds the classification threshold. For each threshold, classification results are re-evaluated in order to depict trade-offs between different metrics. PR curves represent the one-vs-all precision and recall for each class across the range of classification thresholds. ROC curves represent the one-vs-all true positive rates and false positive rates for each class across the range of classification thresholds. Depending on the use case of a given classifier model, the PR and ROC curves can provide valuable insight into the effects of setting a desired level of performance in one metric, such as the one-vs-all true positive rate for a specific class.

Figures 5, 6, and 7 show the PR and ROC curves for each classification method.

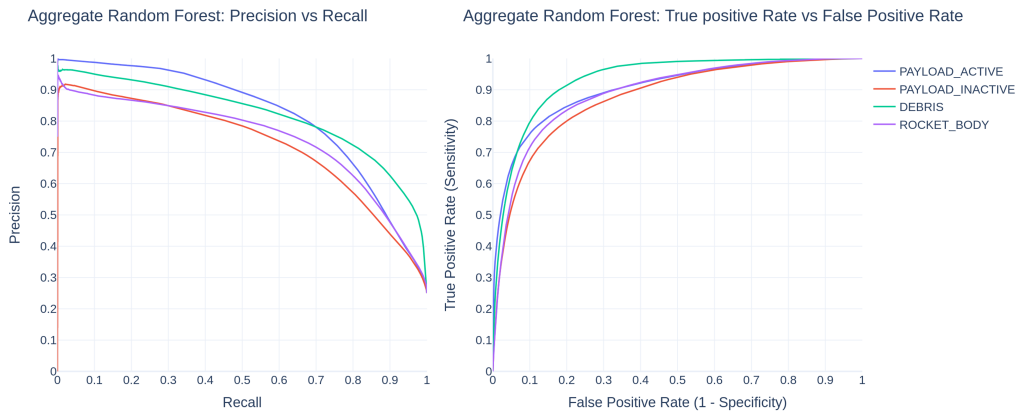

Fig. 5: Precision-Recall and Receiver Operating Characteristic curves for the Aggregate Classifier.

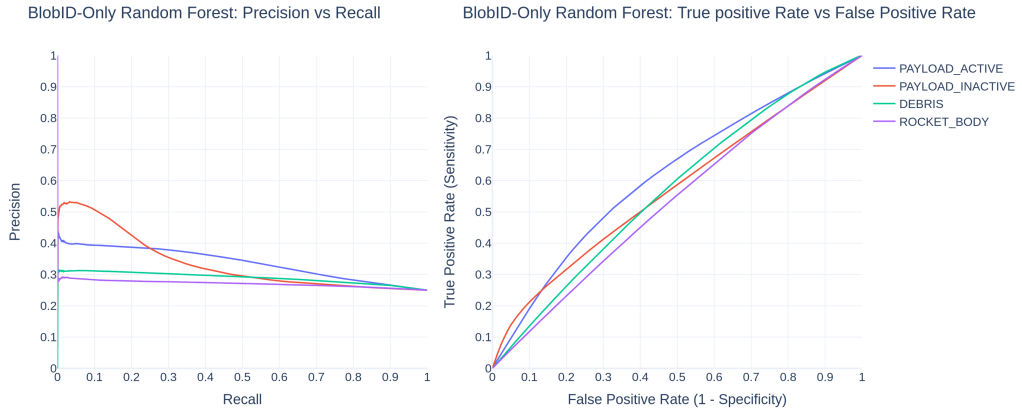

Fig. 6: Precision-Recall and Receiver Operating Characteristic curves for the BlobID Classifier.

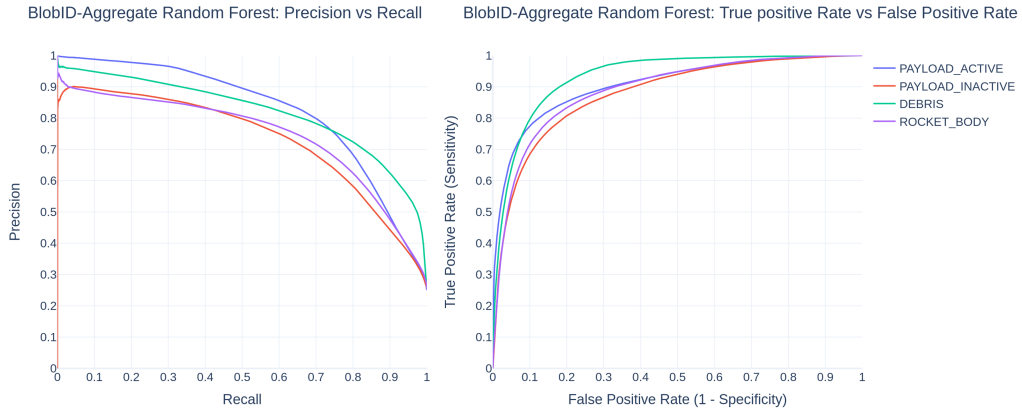

Fig. 7: Precision-Recall and Receiver Operating Characteristic curves for the BlobID-Aggregate Classifier.

As seen in Figures 5, 6, and 7, the two classifiers that use statistical aggregate quantities as input features exhibit more favorable precision/recall and sensitivity/specificity tradeoffs compared to the BlobID-only classifier.

3.3 Confusion Matrices

A confusion matrix is a grid of probabilities that represents the percentages of correct and incorrect classifications for each class. This can yield valuable insight into the specifics of misclassifications in a given model. The values shown in this section are normalized against the actual number of samples in the class, so that the main diagonals are equivalent to the recall for each class. Fig. 8 shows the confusion matrices for the three classifier methods.

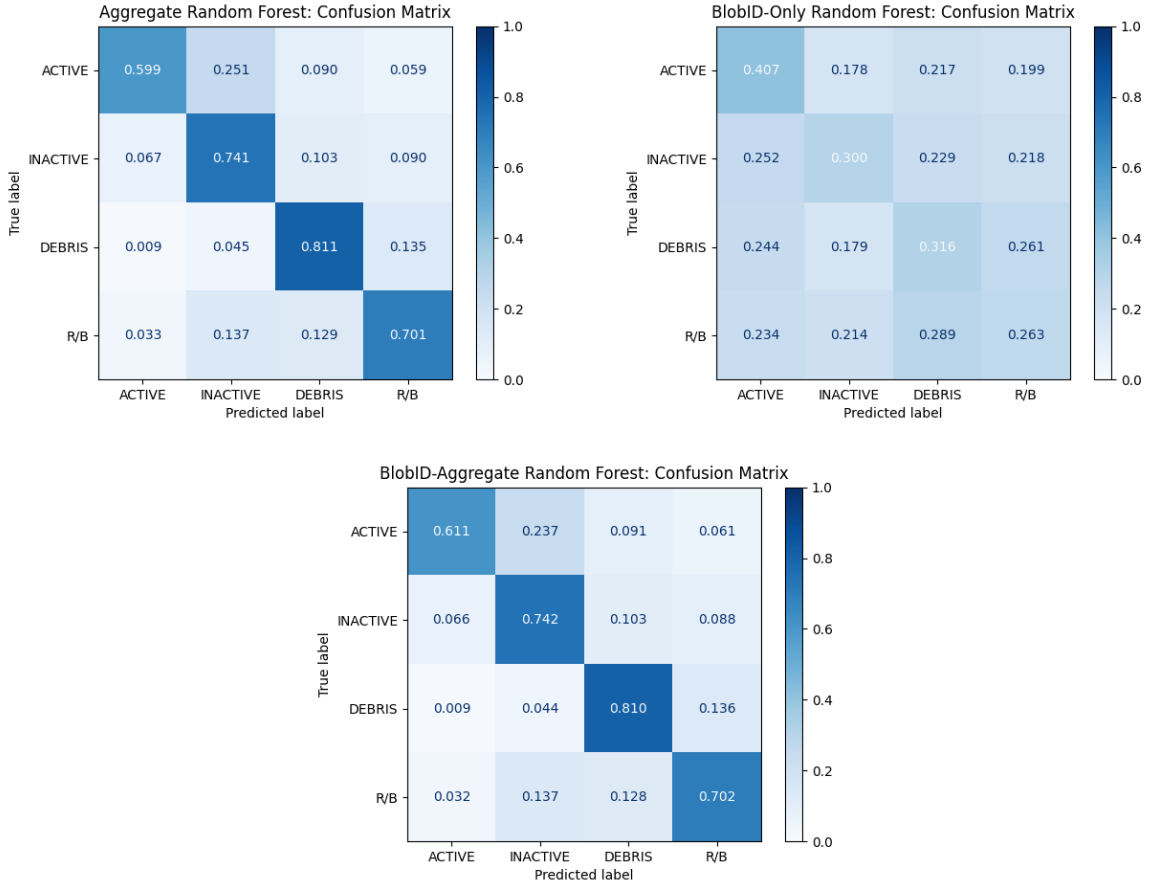

Fig. 8: Confusion matrices for Aggregate (top-left), BlobID-Only (top-right), and BlobID-Aggregate (bottom)

Fig. 8 shows that the BlobID-Aggregate classifier demonstrates about 1.20% higher recall for the active payload class when compared to the Aggregate classifier. This result indicates that the Blob Identification outputs help the Random Forest algorithm resolve edge cases in situations where inactive and active payloads are difficult to distinguish. For both of the classifiers fit using the aggregates of the raw data, the portion of the confusion matrices relating inactive payloads, debris, and rocket body classifications are identical to 2 decimal places, showing that the preprocessing does not provide as much aid in distinguishing these classes.

The confusion matrix for the BlobID-Only classifier provides insight into which classes the Blob Identification methods are most useful for distinguishing class types. The classifier is the most accurate on the active payload class, with a recall of 40.7%. This is an intuitive result because the Blob Identification methods are centered around determining whether or not an object is stabilized, and the active payload is the only class corresponding to stabilized objects. Additionally, the BlobID-Only classifier is able to distinguish debris from inactive payloads relatively well. However,

the recall on rocket bodies is 26.3%, which is only marginally more accurate than random guessing. This method performs poorly when distinguishing between rocket bodies and debris, suggesting that the Blob ID methods output similar results for these two classes.

3.4 Feature Importance

Feature importance is calculated as the mean decrease in impurity for each feature in the Random Forest [30]. Feature importance depicts how much emphasis the Random Forest places on a feature for differentiating between the different classes. A feature importance of zero means that the feature is not used at all in the Random Forest.

Fig. 9 shows the importances of statistical aggregate features used in Method 1.

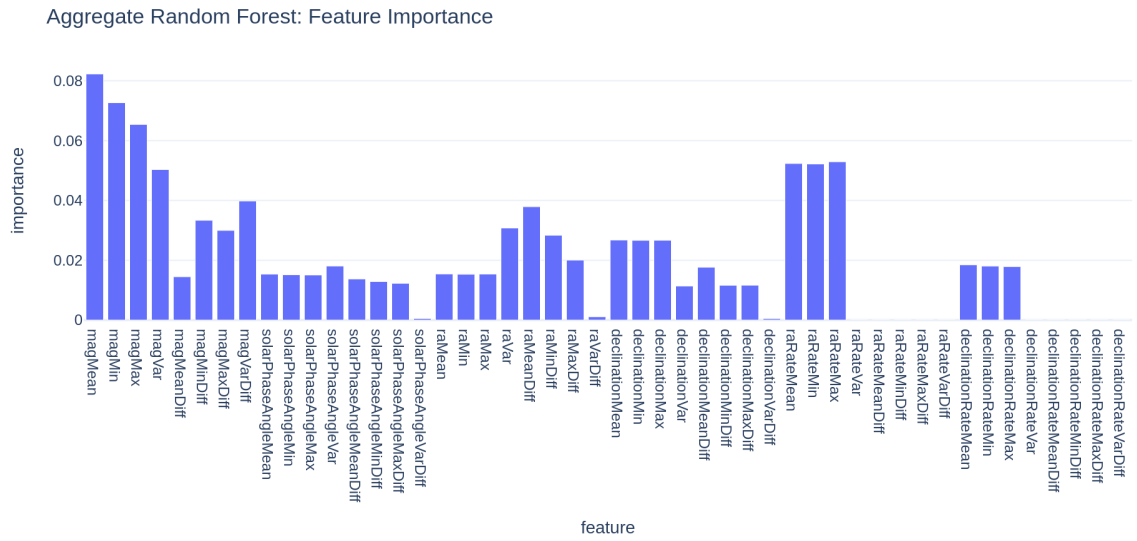

Fig. 9: Feature importances for Aggregate Random Forest classifier.

As seen in Fig. 9, the most important statistical features for class differentiation are the mean, minimum, maximum, and variance of magnitude, and mean, minimum, and maximum of right ascension rate. The high importance associated with right ascension rate is suspected to be due to it being a proxy for orbital altitude. Several features exhibited little to no importance, including the variance of the back-differences of solar phase angle, right ascension, and declination, as well as all features derived from the back-differences of right ascension rate and declination rate. These features could likely be removed without degrading the performance of the classifier.

Fig. 10 shows the importances of the features derived from Blob Identification used in Method 2.

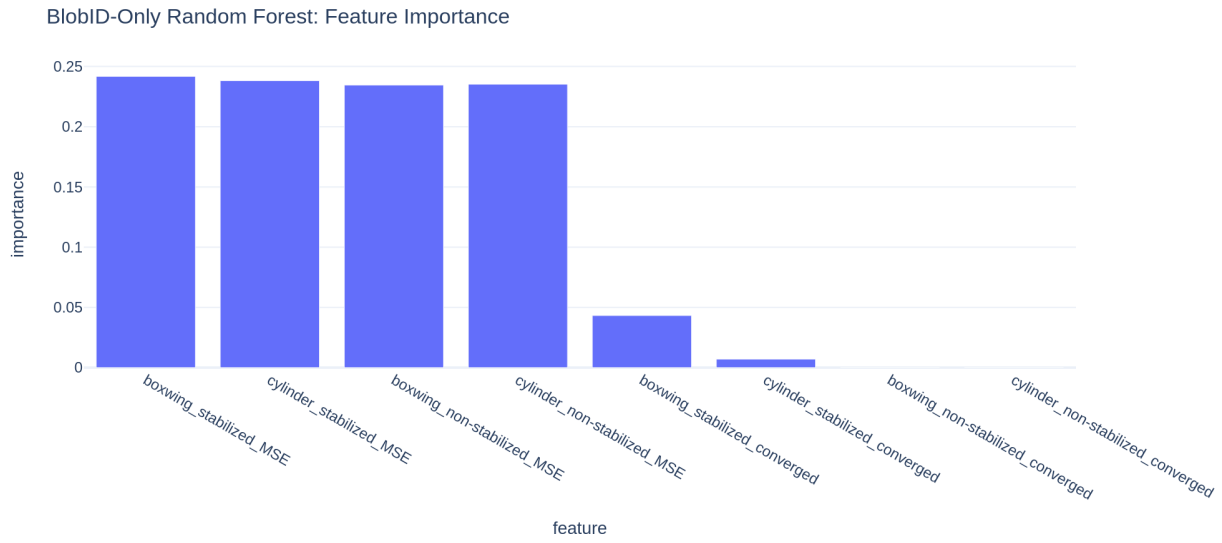

Fig. 10: Feature importances for BlobID-Only Random Forest classifier.

From Fig. 10, it can be seen that the mean squared errors of all four Blob Identification models contribute roughly equally to the class predictions. The convergence flags, on the other hand, contribute much less to the final predictions, with two of these features exhibiting zero importance.

Fig. 11 shows importances of combined statistical aggregate and Blob Identification features used in Method 3.

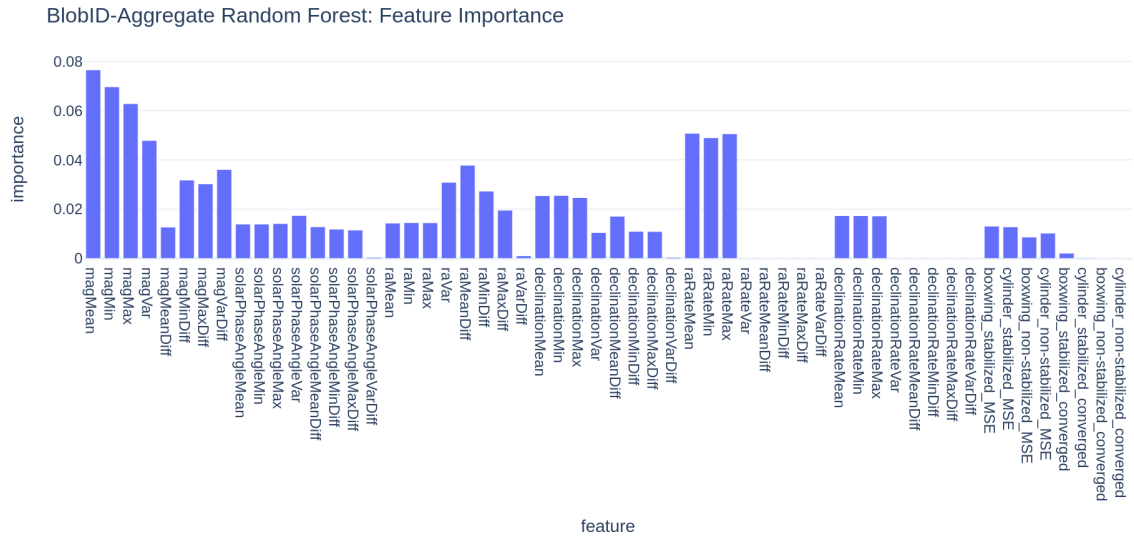

Fig. 11: Feature importances for BlobID-Aggregate Random Forest classifier.

As seen in Fig. 11, the addition of the Blob Identification features did not significantly impact any of the importances of the statistical aggregate features. While the overall importance of the Blob Identification decreased significantly, showing informational overlap with the statistical aggregates, they remained similar to one another. These results further support that most, but not all, of the information provided by Blob Identification for class prediction is captured by the statistical aggregate features. Furthermore, it shows that the Blob Identification features are not highly correlated

to any individual statistical aggregate feature or simple combination of such. Rather, the informational overlap comes from deeper representations of these statistical aggregates that the Random Forest is able to express.

4. CONCLUSIONS

This work demonstrates that object classes, delineated by active payloads, inactive payloads, debris, and rocket bodies, can be predicted from short tracks, composed of 10 to 15 electro-optical observations, with durations of less than 10 seconds. A Random Forest model was used with three different sets of input features: aggregate statistical quantities, data products from partial light curve inversion (also known as Blob Identification, or Blob ID), and a combination of both.

It is found that the Random Forest using statistical aggregate quantities as input features outperforms the model using only Blob Identification data products. There is a marginal performance increase across all measured metrics when all input features are used, showing that there is a small amount of independent predictive power offered by Blob ID when used as an input to the Random Forest. Considering that the Blob ID method is reliant on periodic temporal and phase angle dependent properties of the light curve, it is notable that Blob ID offers any predictive power on tracks as short as those evaluated in this study.

5. ACKNOWLEDGMENTS

The material in this paper is based on work supported by the Department of the Air Force through an Air Force Research Laboratory (AFRL) sponsored Small Business Innovation Research (SBIR) Program Phase I contract.

REFERENCES

[1] M Kaasalainen, Lumme Lamberg, K Lumme, and E Bowell. Interpretation of lightcurves of atmosphereless bodies. i-general theory and new inversion schemes. *Astronomy and Astrophysics*, 259:318–332, 1992.

[2] M. Kaasalainen and J. Torppa. Optimization methods for asteroid lightcurve inversion: I. shape determination. *Icarus*, 153(1):24–36, 2001.

[3] Mikko Kaasalainen and Johanna Torppa. Optimization methods for asteroid lightcurve inversion: I. shape determination. *Icarus*, 153(1):24–36, 2001.

[4] Doyle Hall, John Africano, David Archambeault, Brian Birge, David Witte, and Paul Kervin. Amos observations of nasa's image satellite. In *The 2006 AMOS Technical Conference Proceedings*, pages 10–14, 2006.

[5] Doyle Hall, Brandoch Calef, Keith Knox, Mark Bolden, and Paul Kervin. Separating attitude and shape effects for non-resolved objects. In *The 2007 AMOS Technical Conference Proceedings*, pages 464–475. Maui Economic Development Board, Inc. Kihei, Maui, HI, 2007.

[6] Doyle Hall, Brandoch Calef, Keith Knox, Mark Bolden, and Paul Kervin. Separating attitude and shape effects for non-resolved objects. In *The 2007 AMOS Technical Conference Proceedings*, pages 464–475. Maui Economic Development Board, Inc. Kihei, Maui, HI, 2007.

[7] Siwei Fan and Carolin Frueh. A direct light curve inversion scheme in the presence of measurement noise. *The Journal of the Astronautical Sciences*, 67, 08 2019.

[8] Siwei Fan. *The Light Curve Simulation and Its Inversion Problem for Human-Made Space Objects*. PhD thesis, Purdue University, August 2020.

[9] S Fan and C Frueh. Multi-hypothesis light curve inversion scheme for convex objects with minimal observations. In *Proceedings of the 8th European Conference on Space Debris*, pages 1–7. ESA Space Debris Office, 2021.

[10] Liam Robinson and Carolin Frueh. Light curve inversion for reliable shape reconstruction of human-made space objects. In *Proceedings of the 32nd AIAA/AAS Astrodynamics Specialist Conference*, pages 1–19, 09 2022.

[11] JG Williams. The determination of the orientation of a tumbling cylinder from the shape of the light curve. *OFFICE OF AEROSPACE RESEARCH*, page 31, 1967.

[12] Patrick Koller. Attitude determination of cylindrical rocket bodies from optical light curves. University of Bern, 2016. Bachelor Thesis.

[13] Conor J Benson, Daniel J Scheeres, William H Ryan, and Eileen V Ryan. Cyclic complex spin state evolution of defunct geo satellites. In *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, HI*, 2018.
[14] Alex Burton and Carolin Frueh. Light curve attitude estimation using the viewing sphere. In *Astrodynamics Specialist Conference 2020*, 2020.
[15] A Burton and C Frueh. Two methods for light curve inversion for space object attitude determination. In *8th European Conference on Space Debris*, 2021.
[16] Stephen R. Gagnon and John L. Crassidis. Augmenting light curve based attitude estimation with geometric information. In *AIAA SCITECH 2022 Forum*, 2022.
[17] Mikko Kaasalainen, Johanna Torppa, and Karri Muinonen. Optimization methods for asteroid lightcurve inversion: Ii. the complete inverse problem. *Icarus*, 153(1):37–51, 2001.
[18] Richard Linares, Moriba K Jah, John L Crassidis, and Christopher K Nebelecky. Space object shape characterization and tracking using light curve and angles data. *Journal of Guidance, Control, and Dynamics*, 37(1):13–25, 2014.
[19] Richard Linares, Moriba Jah, and John Crassidis. Inactive space object shape estimation via astrometric and photometric data fusion. *Advances in the Astronautical Sciences*, 143:217–232, 01 2012.
[20] David Gaylor and Jessica Anderson. Use of hierarchical mixtures of experts to detect resident space object attitude. In *Advanced Maui Optical and Space Surveillance Technologies Conference*, volume 1, page 70, 2014.
[21] Richard Linares and John L Crassidis. Space-object shape inversion via adaptive hamiltonian markov chain monte carlo. *Journal of Guidance, Control, and Dynamics*, 41(1):47–58, 2018.
[22] Alexander Agathanggelou, Ryan Houghton, Joshua Collyer, Joshua Davis, and Nicholas Pallecaros. Machine learning for satellite characterisation. *Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference*, 2022.
[23] Ben Lane, Mark Poole, Matt Camp, and Jeremy Murray-Krezan. Using machine learning for advanced anomaly detection and classification. *Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference*, 2016.
[24] Nicholas Perovich, Zachary Folcik, and Rafael Jaimes. Applications of artificial intelligence methods for satellite manuever detection and manuever time estimation. *Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference*, 2022.
[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, page 1097–1105, 2012.
[26] H. Lee, P. Pham, Y. Largman, and A. Y. Ng. Unsupervised feature learning for audio classification using convolutional deep belief networks. *Advances in neural information processing systems*, page 1096–1104, 2009.
[27] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 2015.
[28] Roberto Furfaro, Richard Linares, and Vishnu Reddy. Shape identification of space objects via light curve inversion using deep learning models. *AMOS Technologies Conference, Maui Economic Development Board, Kihei, Maui, HI*, 2019.
[29] T. S. Kelso. Celestrak satellite catalog (satcat) operational status. https://celestrak.org/satcat/status.php.
[30] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
[31] Max Kuhn and Kjell Johnson. *Applied Predictive Modeling*. Springer New York, NY, 2013.
[32] Tin Kam Ho. Random decision forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC*, page 278–282, 1995.
[33] Carolin Frueh. *Identification of Space Debris*. PhD thesis, University of Bern, Switzerland, September 2011.