# ML-Driven Optimal Design of Multispectral Instruments for the Characterization of Resident Space Objects

**Kedar R. Naik, Andrew I. Wernersbach, Alexandra L. Robinson, Michelle F. Nilson, Gary D. Wiemokly, Matthew N. Tooth, Raymond H. Wright**
*BAE Systems Inc., Space & Mission Systems*

## ABSTRACT

The characterization of resident space objects (RSOs) has long been an important aspect of space domain awareness (SDA). Recently, spectral imagery has emerged as a viable sensing modality with the potential to enhance existing RSO-characterization capabilities, especially when used alongside machine-learning (ML) models. Deploying a spaceborne hyperspectral sensor, however, can increase size, weight, power, and cost (SWaP-C). On the other hand, multispectral sensors, which collect significantly fewer bands than hyperspectral sensors, are less complex to develop and deploy. This paper presents an optimization procedure that finds the minimal number of multispectral bands required for training high-performance ML classifiers. The optimization procedure was used to perform a design study for a notional SDA mission responsible for classifying three types of RSOs: active payloads, rocket bodies, and space debris. Simulated hyperspectral signatures of each class were generated and used to initialize the optimal-design procedure. These hyperspectral signatures consisted of 1,574 spectral bands. The procedure was able to design an optimized multispectral system with only five bands. Despite the 99.7% reduction in the number of spectral bands, the overall performance of the ML-driven RSO-classification models only decreased by 1.9%. This result suggests that future SDA missions may be able to replace hyperspectral sensors with bespoke multispectral ones while still meeting requirements related to RSO-characterization capability.

## 1. INTRODUCTION

The ability to quickly and accurately characterize resident space objects (RSOs) is a critical component of any space-based space domain awareness (SDA) mission. Proper characterization entails processing raw observables in such a way that facilitates a variety of downstream tasks, e.g., space-object classification, track association, and track initiation [1]. The success of these downstream tasks is predicated on two considerations that have historically been treated separately during the design process: (1) the information content of the raw observables collected by on-board sensors and (2) the predictive power of the algorithms available to process those observables. In recent years, the performance of those processing algorithms has been significantly improved with the advent of easily deployable machine-learning (ML) models [2]. Consequently, there has been growing interest in exploring how the design of on-board sensors themselves might be optimized given that the raw observables can now be processed using ML. The present study seeks to contribute to this novel paradigm for designing imaging systems for SDA missions. The optimization procedure described herein allows the development of electro-optical (EO) sensors and the training of ML-driven processing models to inform one another. This coupled approach to the design of remote-sensing systems promises to minimize overall mission costs while maximizing overall mission performance.

### 1.1 An SDA Mission: RSO Classification

Consider a *generic space-based SDA mission* in geostationary orbit (GEO), which seeks to classify observed RSOs into one of the three most commonly observed categories, namely: payloads, rocket bodies, and debris [3]. (Here, let "rocket bodies" serve as a shorthand for any core stages, first stages, or upper stages that have been abandoned in orbit—as opposed to having been brought back to Earth in a controlled manner [4].[1]) Assume that observations will be made using a passive EO sensor producing unresolved imagery, wherein targets of interest typically occupy less than

[1] It is estimated that 60% of launches result in at least one rocket body being abandoned in orbit. The risks to space missions of all types posed by these derelict rocket bodies is of growing concern [5].

a single pixel. Furthermore, assume there exist ML models trained to classify each pixel of an unresolved image.[2] It follows that, generally speaking, the accuracy of those ML classifiers will increase as the information content of each pixel increases.

### 1.2 EO Sensors for SDA

Increasing pixelwise information content is a perennial challenge for any space-based remote-sensing mission. The large observer distances inherent to SDA missions mean that there is typically a paucity of photons available for image generation. The photon-starved environment necessarily introduces a trade-off between spatial resolution and instantaneous spectral refinement. If the number of photons available to given pixel must be distributed amongst multiple wavelength bins (i.e., spectral bands), the measured dimensions of that pixel in space must increase. Therefore, maximizing spatial resolution in low-light environments means minimizing the number of simultaneously collected spectral bands.

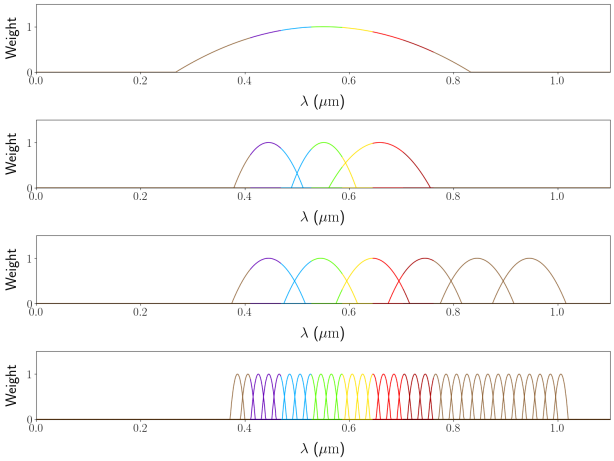



Fig. 1: Notional spectral-response functions of the various types of passive EO sensors. The topmost plot shows the bandpass of a panchromatic sensor; the second plot show the three bands of a traditional RGB or BVR sensor; the third plot shows a simple six-band multispectral sensor, covering the visible portion of the spectrum and extending into the near-IR; the bottommost plot shows the many narrow bands associated with a hyperspectral imager.

#### 1.2.1 Panchromatic

Historically, this has been the approach used for SDA missions. *Panchromatic* sensors collect only a single channel with a broad bandpass, often spanning the entirety of the visible spectrum (from approximately 380 nm to 750 nm) and possibly even extending out into the near-infrared (IR) and ultraviolet (UV) regions. Fig. 1 visualizes the notional bandpass of such an imager. Although panchromatic sensors make the best use of the available light, produce imagery with high spatial resolution, and are relatively straightforward to manufacture, the pixelwise information content they deliver is relatively low. Indeed, all panchromatic imagery is, by definition, colorless. As mentioned above, RSO targets of interest typically fill less than a single pixel of any image collected during an SDA mission. As a result, downstream characterization tasks, such as classification, identification, and tracking must be done using only a single grayscale value—a challenging prospect.

#### 1.2.2 BVR Composite

To imbue each pixel with slightly more information, tri-color composite imagery, wherein each pixel comprises a *Blue-Visual-Red (BVR) light curve*, is often used in place of panchromatic imagery [6]. A BVR light curve is the three-band spectrum produced by passing light through three specifically designed optical filters. The bandpass ranges of these filters roughly correspond to the familiar red-green-blue (RGB) bands used in consumer electronics and digital-imaging applications.[3] Note that the B, V, and, R bands are each collected using *separate* optical filters. This

[2] Readers with a background in modern ML techniques might wonder why models would be trained at the level of the pixel instead of at the level of the image, especially considering the wide variety of deep-learning (DL) architectures that have proven especially performant in the realm of image processing over the past decade (e.g, convolutional neural networks (CNN) and vision transformers). Firstly, the rationale for using pixelwise models flows from the nature of unresolved imagery. During a typical SDA mission, targets of interest seldom occupy more than a single pixel. Consequently, there is limited utility in a neural network's ability to learn subtle spatial relationships. (A couple of obvious exceptions here involve observing satellites swarms engaged in formation flying or observing clouds of space debris. That being said, it is well to mention that it is possible to characterize swarms and clouds, even when using ML models trained at the level of the pixel. The pixelwise results simply need to be post-processed in a way that accounts for spatial information, e.g., via the use of an unsupervised clustering algorithm.) Secondly, onboard-storage constraints may prevent a complete two-dimensional image from being constructed, precluding the use of standard image-processing techniques. Instead, rows of pixels must be processed as they are captured.

[3] The B filter typically captures wavelengths between 400 nm and 500 nm; the V or G filter usually covers the range between 500 nm and 600 nm; and the R filter generally spans the interval between 600 nm and 750 nm. In some cases, the BVR light curve might be augmented to include one or two additional bands covering small portions of the UV and IR regions of the spectrum. In such situations, the initialisms describing the resulting pixelwise spectra are extended accordingly, e.g., "UBVRI light curves."

means that the three color channels comprising the composite image cannot be collected simultaneously. Instead, they are collected in quick succession, allowing each band to be informed by all received photons. A notional sketch showing the spectral response of a BVR imager appears in Fig. 1. Widely used for astronomical applications, where the temporal evolution of light curves is studied for scientific purposes, BVR composites have also been used for RSO classification [7]. Although a BVR imager does not record information from as many wavelengths as does a panchromatic sensor, the varying levels of intensity captured in each of the three wavelength bins allow distinguishing patterns to emerge in both the temporal and spectral dimensions.

#### 1.2.3 Multispectral

The third plot in Fig, 1 depicts the aggregate spectral response associated with a six-band *multispectral* imager. Multispectral imaging takes the idea of filter-based sensing and extends it to include many more spectral bands, often extending well into the IR regions of the spectrum. Multispectral imagers are typically composed of approximately a dozen bands or less. Indeed, a BVR imager is simply a specifically tuned three-band multispectral sensor. The additional bands of a multispectral sensor, however, allow a more detailed spectrum to be measured in each pixel. Having access to such a discrete spectrum can allow downstream tasks, such as RSO classification and material identification to be performed with greater accuracy. This is especially true if the set of multispectral filters has been intentionally selected to reflect wavelengths with the most discriminating power.

#### 1.2.4 Hyperspectral

*Hyperspectral* imaging, also known as imaging spectroscopy, collects a long and nearly continuous spectrum in each pixel. Using either a grating or a prism, the incoming light is dispersed into its many constituent spectral bands [8]. The spectral resolution in a hyperspectral image is extremely fine—often on the order of tens of nanometers. The wavelengths spanned can start in the near-UV regions and continue deep into the long-wave IR (approximately 10 µm). For the sake of explanation, the last plot in Fig. 1 shows the spectral response of a hyperspectral system that begins in the UV and goes into the short-wave IR. The fact that each pixel houses a nearly continuous spectral signature means that the pixelwise information content delivered by a hyperspectral imager can be extremely high. Consequently, the prospect of using hyperspectral imagery for SDA applications has generated considerable interest [9]. Notably, over the past couple decades, there have been several studies in which hyperspectral signatures of space objects have been measured in laboratory environments [10][11], from ground-based telescopes [12][13][14], and even on the International Space Station [15].

### 1.3 Challenges with Hyperspectral Sensing

Collecting hyperspectral imagery requires the use of a spectrometer. Fielding an imaging spectrometer for either an airborne or spaceborne remote-sensing campaign can be notoriously expensive. Although there do exist relatively affordable, commercial-off-the-shelf (COTS) spectrometers, these instruments are generally not suitable for higher-altitude airborne missions or spaceborne missions. One reason for this is the extreme thermal environment, which can cause optical misalignment. Correcting for this misalignment requires the ability to perform on-orbit spectral calibration or the integration of a thermal-control system—both of which can add significant cost and complexity to the overall mission. Moreover, for spaceborne missions specifically, further precautions must be taken to ensure that the spectrometer can survive the vibrational and gravitational loads associated with being launched into orbit. COTS spectrometers are simply not designed with such degrees of robustness in mind. Building and integrating tailor-made instruments capable of surviving launch and operating in space environments adds to the overall size, weight, power, and cost (SWaP-C) of the space vehicle.

Additional challenges associated with spaceborne sensors flow from the concept of operations by which collection is meant to be performed. Because the dispersion element of a spectrometer separates incoming light into its many constituent wavelengths, the number of photons representing any particular wavelength is limited. Consequently, in order to achieve useful signal-to-noise ratios during collection, long dwell times are required. Collection time is further increased due to the spatial-scanning approach—either using a pushbroom or whiskbroom concept—adopted by the imager in question [16]. There are practical implications for an SDA mission here. First, scanning mechanisms (e.g., fast steering mirrors) must be procured and integrated such that the instrument's field of view can be swept, line-by-line, across the the region of interest. In addition to extending hardware-development timelines, these scanning

mechanisms add to the overall SWaP-C and operational complexity of the payload. Second, and perhaps more salient from the perspective of mission operations, is the fact that longer per-image collection times necessarily delay the delivery of mission products to stakeholders on the ground. That is, hyperspectral sensing for space-based SDA brings with it the opportunity costs associated with not having timely collection on all regions of interest.[4]

Even if all desired collections are made in a timely manner, however, the images themselves may be difficult to store. Each hyperspectral image is a three-dimensional array, or *data cube*, representing two spatial dimensions and one spectral dimension. The spectral dimension can comprise hundreds of wavelength bins. This means that a single hyperspectral data cube can easily require the same amount of onboard storage as hundreds of panchromatic, tri-color composite, or multispectral images. This fact present mission architects with a trade-off: either decrease the collection capacity between ground passes or increase the amount of onboard storage—implicitly accepting a commensurate increase in SWaP-C.

### 1.4 Advantages of Multispectral Sensing

A bespoke multispectral system has the potential to circumvent many of the complications highlighted above. Since each filter in a multispectral system has access to all the photons in an incoming beam of light, the signal-to-noise ratios are inherently much higher, which in turn allows dwell times to be shorter. Similarly, the use of filters obviates the need for scanning, since the field of view is typically comparable to the region of interest. This larger field of view also eliminates the aforementioned collection-timing constraints. Production costs and operational complexity are also reduced with a multispectral system; the process by which optical filters are manufactured is mature, the filters themselves are less sensitive to on-orbit thermal fluctuations, and high-performance filter wheels are readily purchased from COTS vendors. Moreover, since a multispectral sensor records data at far fewer spectral bands than does a spectrometer—sometimes representing over an order of magnitude difference—the onboard-storage requirements for multispectral sensing are considerably less than those for hyperspectral sensing. This, in turn, alleviates many of the concerns associated with supporting extended collection sequences. Taken together, multispectral imaging offers several advantages over hyperspectral imaging, especially from the perspective of SWaP-C budgets and collection-timing constraints.

### 1.5 Goals of the Present Study

Despite the challenges associated with hyperspectral sensing—and the potential benefits offered by multispectral sensing—it is still the case that hyperspectral instruments capture nearly continuous, high-resolution spectral signatures in each spatial pixel. As suggested in Sec. 1.1, the performance of ML models increases with pixelwise information content. A natural conclusion here is that the best ML models will be those trained on hyperspectral data. The implicit corollary is that models trained on multispectral data will be much less performant.

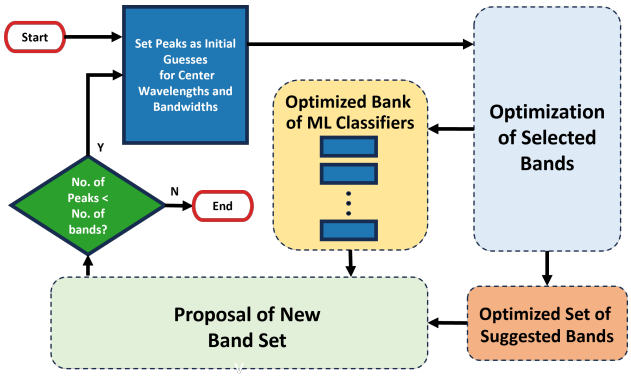



Fig. 2: A summary of the optimal band-design procedure. At each iteration, an optimal set of multispectral bands is produced, along with a bank of tuned ML classifiers.

The findings and the results of the present study suggest that the previous statement is untrue. Although hyperspectral imagery provides considerably more *pixelwise data* than does multispectral imagery, it does not necessarily yield more *pixelwise information content*. This paper posits that for any given SDA mission (e.g., RSO classification in GEO), highly performant ML models can be trained using data from only a small set of intentionally selected bandpass regions. That is, the performance of ML models trained on hyperspectral signatures will be nearly indistinguishable from the performance of ML models trained on multispectral signatures defined at an *optimal set of bands*. In support of this proposition, the authors present an *optimal band-design procedure* that accepts from the user a set of hyperspectral signatures representing entities to be classified (e.g., RSOs) and returns:

[4]One way to overcome these timing constraints would be to launch multiple space vehicles, thereby allowing multiple hyperspectral sensors to perform collects within a region of interest concurrently. Naturally, this solution would increase overall mission costs severalfold, requiring advanced data-fusion algorithms, novel collection-orchestration protocols, and enhanced ground-operations support.

- The minimal set of multispectral bands required to classify those entities
- A bank of ML models trained to classify each entity using spectral curves defined at that minimal set of bands.

The flowchart seen in Fig. 2 adumbrates the iterative optimization procedure. Starting with hyperspectral signatures, each iteration of the process proceeds by (a) optimizing the given set of bands and then (b) proposing a new, smaller set of bands. This process will be expounded upon in the following section.

### 1.6 Outline

The remainder of this paper is organized as follows. Sec. 2 provides an in-depth explanation of each component of the optimal band-design procedure. This section also includes a discussion of how hyperspectral signatures were generated for initializing the optimization process. Sec. 3 shares the results of a notional design study, wherein the optimal set of bands is found for the RSO-classification mission introduced in Sec. 1.1. The paper closes with Sec. 4, which offers some concluding remarks.

## 2. METHOD

As alluded to in Fig. 2, the optimal band-design procedure comprises several different algorithmic components, all of which work in tandem over the course of a given iteration. The algorithmic components, themselves, are responsible for a variety of tasks, from ML training to Bayesian optimization to model explanation. This section presents the details of each component and also describes how they fit together.

### 2.1 Simulation of Hyperspectral Signatures

The optimal band-design procedure begins by ingesting hyperspectral signatures representing the desired classes. The procedure places no constraints on the nature of these signatures. They can span any range of the electromagnetic spectrum and they can be recorded at any spectral resolution. Likewise, there is no restriction how the signatures are sourced; they may be generated using simulation tools, collated from actual measurements, or represent some combination of the two. These hyperspectral signatures are used directly only during the initial iteration of the design procedure. At each subsequent iteration, they are spectrally downsampled to the current set of proposed bands. For the present study, hyperspectral signatures representing the three classes of RSO—i.e., active satellite payloads, abandoned rocket bodies, and space debris—were simulated using a physics-based software package.

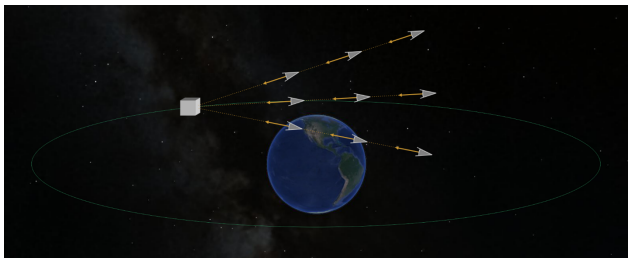
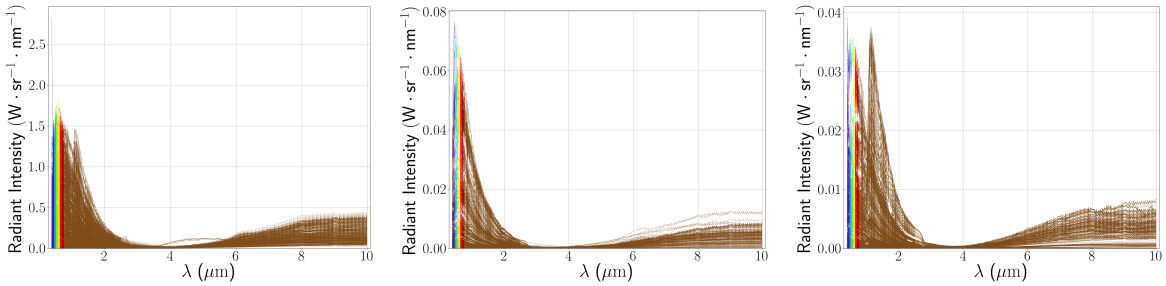


Fig. 3: A visualization of the nine observers that maintain range and attitude with respect to the RSO target as it moves through its orbit in GEO. N.B. Sizes and angles are not to scale.

Physics-based simulation of hyperspectral signatures of RSOs has been performed by several previous authors, using a variety of software packages. Recently, for example, Najera et al.[17] simulated ground-based spectral measurements using the popular DIRSIG package [18]. Last year, Hernandez et al. [19] shared the output of their newly developed first-principles hyperspectral scene-generation tool, which models low-level properties related to a sensor's optics and detector. Earlier this year, Coiro et al. [20] presented the SIRIUS tool, which simulates space-based hyperspectral imagery of RSOs at high spatial resolution while modeling several physical details, including the spectral bidirectional reflectance distribution functions of various materials, the isolated rotation of the different parts of an RSO (e.g., the solar panels), and the radiative effects of other celestial bodies.

#### 2.1.1 Data Generation

For the present study, simulated hyperspectral signatures were generated in order to model the generic SDA mission described in Sec. 1.1. For each of the three classes of RSO, a variety of class-specific surface compositions were compiled and applied to a representative geometric model. Each surface composition was observed by a sensor model from nine different viewing geometries and 180 different solar phase angles in GEO. Simulations were performed over a 24-hour period, so that the full range of solar phase angles could be captured.




(a) Active payload (b) Abandoned rocket body (c) Space debris

Fig. 4: The three subfigures plot approximately 50% of the hyperspectral signatures generated for each of the three RSO classes of interest.

**RSO-Specific Surface Compositions** To model the *active payload* class, six unique surface combinations were generated, involving representative metals, plastics, and glass. For the *abandoned rocket body* class, again six surface compositions were modeled, however in this case, the constituent materials were representative metals and paints. To represent the *space debris* class, seven of the materials used in modeling the first two classes were selected.

**Sensor Model** An EO ultra-wideband sensor with low noise and high quantum efficiency, capable of detecting light across a wide range of wavelengths from the visible through the long-wave IR, was modeled for the present study. The sensor parameters used are presented in Table 1. N.B. In the table, FOV denotes field of view and IFOV denotes instantaneous field of view.

Table 1: Parameters used to define the sensor model.

| Parameter | Value |
|---|---|
| Pixel IFOV | 2.9 µrad |
| Simulated FOV (16×16 pixels) | 0.003° |
| System FOV (2000×2000 pixels) | 0.3° |
| Focal Length | 2.1 m |
| Aperture Diameter | 20 cm |
| F-number | 10.5 |
| Wavelength Range | 0.4 µm – 10 µm |
| Channel Type | Flat Top |
| Channel Spacing | 50 nm |
| Channel Width | 50 nm |
| Quantum Efficiency | 0.8 |
| Electron Well Depth | 100,000 $e^-$ |
| Digitization | 14 bits |
| Digitization Gain | 6.10 $e^-$/DN |
| Dark Current | 10 $e^-$/s |
| Read Noise | 10 $e^-$ |
| Integration Time | 0.1 s |

**Observer Locations** For each modeled surface composition, nine observers, i.e., nine copies of the sensor model, were used to simulate the collection radiance profiles from nine distinct positions, defined relative to the RSO target. Three lines of sight were defined from the target. One lay ahead of the target, along its velocity vector; the second was placed above the orbital plane, $+30°$ from the velocity vector; and the third was placed below the orbital plane, $-30°$ from the velocity vector. Three observers were positioned along each line of sight at fixed ranges of 6 Mm, 8 Mm, and 10 Mm. The observer positions and attitudes were fixed relative to the target as the target progressed through its orbit. As such, each observer's view of the target was constant while the solar phase angle varied. Fig. 3 notionally shows the three lines of sight and the three observers along each line of sight at distinct ranges.

After removing corrupted data, the process described above yielded a total of 2,156 hyperspectral signatures. Specifically, there were 967 spectra representing active payloads, 704 representing abandoned rocket bodies, and 485 representing debris particles. A representative sampling of the spectral signatures corresponding to each of the three classes of RSO appear in Fig. 4. In those plots, visible regions of the spectrum are colored as they would appear to the human eye, whereas IR regions are colored maroon.

### 2.1.2 Data Partitioning

As will be discussed further in subsequent sections, each iteration of the optimal band-design procedure comprises three sequentially linked algorithmic components, viz. the training and tuning of ML models, the optimization of a given set of bands, and the proposal of a novel set of bands. In order to avoid leakage between data samples used for training, tuning, and evaluation, the overall dataset of 2,156 spectral signatures was partitioned several times. These partitions are depicted schematically in Fig. 5. As sketched in the figure, the five-way partition is performed in two steps. First, the data samples are randomly shuffled. Then, a stratified train-validation-test split is performed at 70%, 10%, and 20%, respectively. At the start of each trial of the *band-optimization module* (colored light blue in Fig. 2), the 70% partition is, itself, partitioned using the same fractions. The result is five distinct datasets, namely: a single training set ("Train 2"), two validation sets ("Validation 1" and "Validation 2"), and two test sets ("Test 1" and "Test 2"). The precise manner in which these partitions are employed during the overall optimal band-design procedure is detailed in the next section.

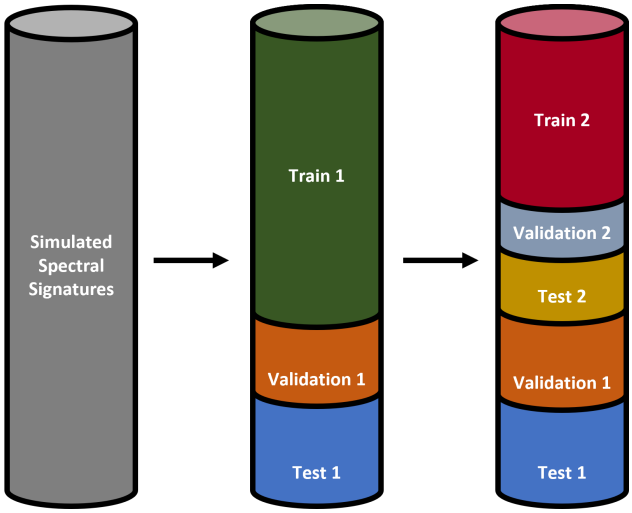



Fig. 5: In order avoid data leakage and unbiased evaluation results, the overall dataset of simulated spectra is partitioned into five separate datasets. N.B. "Train 2," "Validation 2," and "Test 2" are regenerated at the start of each trial of the band-set optimization module.

## 2.2 Optimal Band-Design Procedure

As sketched in Fig. 2, the band-design procedure developed for the present study is iterative in nature. At each iteration, the current set of bands is adjusted in such a way that maximizes the performance of the ML classifiers. Explainable-artificial-intelligence (XAI) techniques are then used to determine which bands from the current band set were most responsible for improving the accuracy of the ML models. Finally, a new band set is generated based off those XAI results. If the newly proposed band set is smaller than the current band set, the next iteration commences.

The remainder of this subsection seeks to convey a general overview of the procedure before examining its inner workings. To that end, the flow of information through each iteration is first summarized at a high-level. Then, each of various algorithmic components is treated in detail. Prior to introducing the iterative flow, however, some relevant work by other authors is summarized in order to provide the reader with additional context and background.

**Previous Work** The iterative optimization procedure of this paper has been developed to help payload engineers and SDA mission architects find SWaP-C savings during the design process without sacrificing overall mission performance. Although previous work on optimal band design is limited, there is a small handful of recent studies that deserves mention here. In 2016, Zhao et al. [21] took ground-based, ten-band multispectral measurements of satellites in GEO. They showed that four different types of satellites could be distinguished using correlation coefficients. They then attempted their analysis again with only five bands, excluding data from every other band. They found that some ability to distinguish the satellites remained.[5] More recently, Hernandez et al. [19] developed a high-fidelity, spectral scene-simulation framework for SDA missions. The framework include the ability to train DL classifiers on the synthetic scenes, thereby allowing ML/DL-informed trades to be performed during the sensor-design process. Earlier this year, Ozge and Atik [22] used popular XAI techniques to perform band selection (not band design) on existing hyperspectral images. Although the literature is replete with such band-selection studies, this one is notable for its use of modern XAI methods. Also this year, Naik et al. [23] presented a precursor to the present study, wherein the optimal band-design procedure was applied to a notional geospatial-intelligence mission.

[5] N.B. This brute-force elimination of bands is quite distinct from the present study, wherein the bands are intentionally designed.

### 2.2.1 Initial Iteration

As tends to the be case with most iterative algorithms, the first iteration here stands apart from all subsequent iterations. The initial iteration of the optimal band-design procedure processes the original set of spectra, recorded at hyperspectral resolution. As the objective of this procedure is to *reduce* the number of bands required for hyperspectral sensing, there is little utility in optimizing the center wavelengths and bandwidths of the original band set. So, for the first iteration, the *band-optimization module* may be simplified significantly. Instead of repeatedly adjusting the bands, downsampling the training data, and retraining the ML models, the ML models are only trained once. Dataset "Train 1" (see Fig. 5) is used to provide the training, validation, and test set for each classifier. The aggregate, multiclass performance of all the models is evaluated using dataset "Test 1." These models are then used by the *band-proposal module* to identify important regions of the spectrum and propose a new set of bands. The processing steps performed during the initial iteration are summarized in Fig. 6.

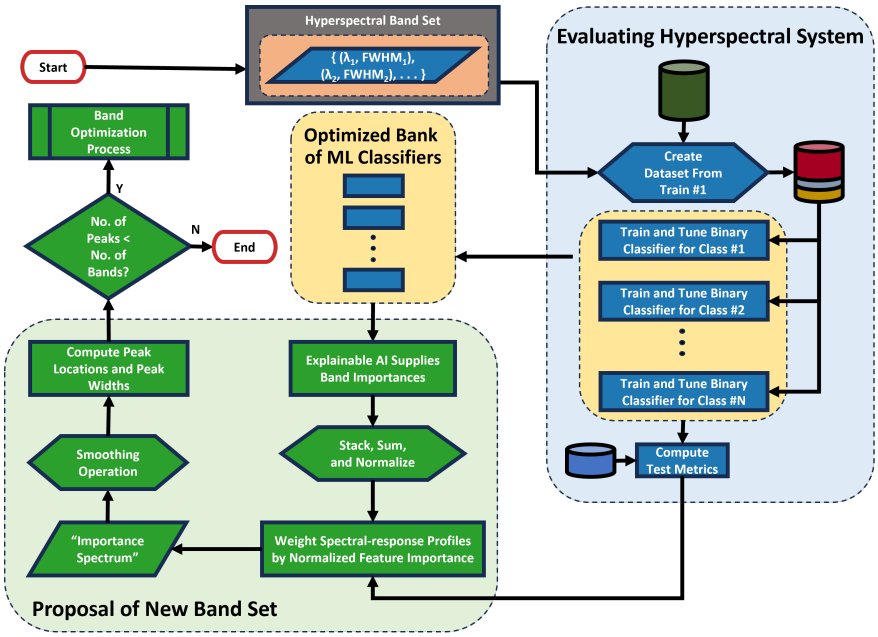



Fig. 6: The flowchart depicts the initial iteration of the optimal band-design procedure. Since the procedure begins by treating the hyperspectral signatures, optimal adjustment of the bands is not performed. The datasets used for the various steps are indicated using cylinders, whose colors match those seen Fig. 5.

### 2.2.2 Subsequent Iterations

At each iteration after the first, the newly proposed set of bands is first optimized before high-importance regions of the spectrum are identified. The overall flow of these subsequent iterations is presented in Fig. 7. The band-optimization module now involves multiple steps:

1. Downsampling the training data (dataset "Train 1")
2. Partitioning "Train 1" into the "Train 2," "Validation 2," and "Test 2" datasets
3. Using those partitions to train and tune binary classifiers for each class of RSO
4. Measuring the multiclass performance of those classifiers (using the "Validation 1" dataset)
5. Based on that performance, using an optimizer to adjust the center wavelengths and bandwidths of each band.

It is worth noting that the band-optimization module is, itself, an iterative process. To avoid confusion, each iteration of this module if referred to as a *trial*. The number of trials performed for each set of proposed bands is specified by the user.

Once the desired number of optimizer trials has been performed, the result is an *optimized set of bands* and a corresponding *bank of optimized ML classifiers*. Note that each classifier in that bank is trained to classify spectra defined at the optimized set of bands. These results are then fed to the module responsible for proposing the next set of bands. Like the band-optimization module, the band-proposal module comprises multiple steps:

1. Using XAI techniques, compute the importance of each band for each of the trained ML classifiers
2. Sum these importance values for all of the classifiers, yielding the importance of each band to the overall mission
3. Normalize these *mission-level importance values* across the the wavelengths such that they sum to unity
4. Multiply each spectral-response function defining the current band set by the corresponding mission-level importance value
5. Construct the *importance spectrum* for the current band set by summing together the importance-weighted spectral-response functions, yielding a continuous curve across all the wavelengths
6. Apply a smoothing operation to the importance spectrum
7. Locate each peak in the importance spectrum, noting its position along the wavelength axis and corresponding width
8. Propose a new set of bands in which the center wavelengths are defined at the peak positions and the bandwidths are defined by the corresponding peak widths.

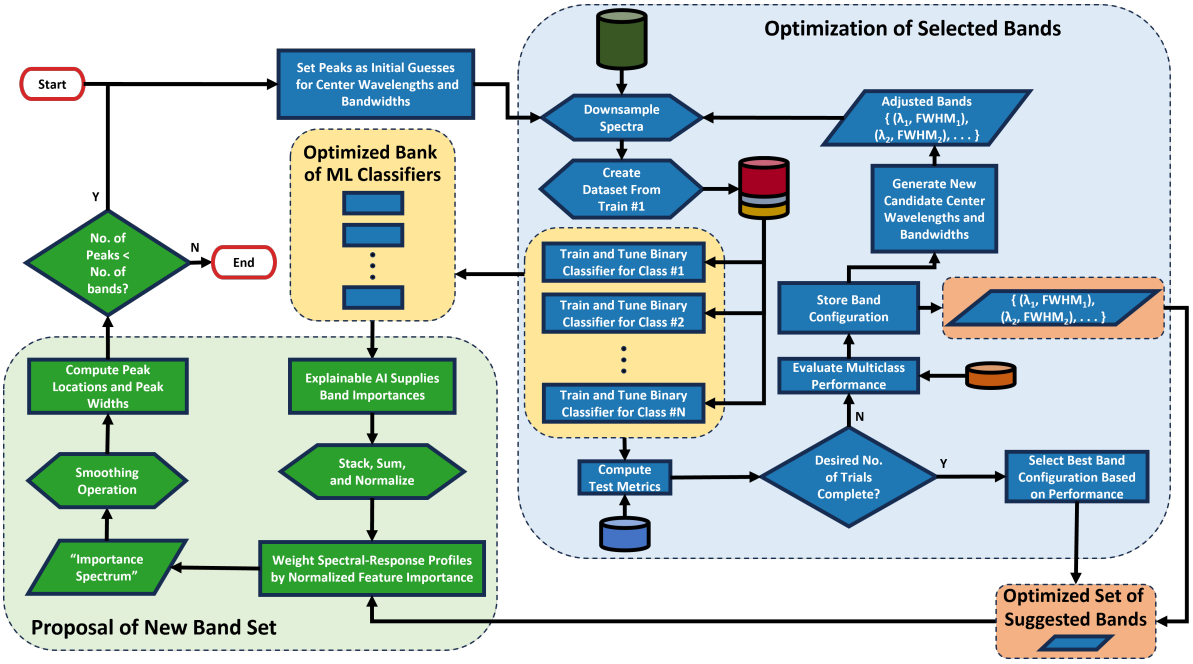



Fig. 7: The flowchart depicts a typical (i.e., non-initial) iteration of the optimal band-design procedure. An optimizer adjusts the current band centers and bandwidths by attempting to maximizing the resulting performance of the ML models. A new band set is then proposed based on the importance assigned to each band by XAI techniques. The datasets used for the various steps are indicated using cylinders, whose colors match those seen Fig. 5.

### 2.2.3 Stopping Criterion

The output of the band-proposal module is a new set of importance-informed bands. If the number of bands in the newly proposed set is smaller than that of the current band set, then the procedure continues on to the next iteration. If, however, the number of bands is greater than or equal to the current number of bands, the optimal band-design procedure is halted and the current band set is returned.

Note that convergence in the number of bands is not the only available stopping criterion. Perhaps the user is more interested in maintaining the performance of the ML classifiers than completely minimizing the total number of bands. In this case, the procedure ends when the multiclass performance of the optimized bank of ML classifiers falls below some desired threshold.

### 2.2.4 Training and Tuning of ML Classifiers

Training and tuning ML classifiers is an important component of the band-optimization module. During the course of a band-design study, this module trains classifiers on either hyperspectral or multispectral data. In this work, the same underlying ML algorithm is used, irrespective of the number of bands: gradient boosting.

A review of the literature reveals that a variety of different ML/DL algorithms have been applied to the problem of RSO characterization. Some recent studies are mentioned here for completeness. A good review of work in this area has been compiled by Tsaprailis et al. [2]. Recently, Phelps et al. [24] used CNNs to regress RSO orientation from spectral images. In a related study, Gazak et al. [25] used the ResNet architecture to classify RSOs. Li et al. [26] made use of three-dimensional CNNs to identify RSO materials in hyperspectral imagery. Earlier this year, Vasile et al. [27] used a multilayer perceptron (MLP) to regress material compositions from single pixels of hyperspectral imagery. A few months later, Cedillo et al. [28] presented work wherein an MLP was used in conjunction with a convolutional autoencoder to classify satellites using ground-based hyperspectral imagery.[6]

**Binary Classification** For each of the three RSO classes of interest, binary-classification models were trained in a one-versus-rest manner. A bank of binary classifiers was used instead of a single multiclass classifier because of the potential versatility of the former during an actual SDA mission. Consider, for instance, a situation where the ability to detect a new class of RSO has suddenly become important to the mission. Instead of having to retrain an entire multiclass-classification model, only a single binary classifier needs to be trained and added to the existing bank of models. Similarly, if a certain type of RSO is no longer of interest to the mission, the corresponding binary classifier can simply be removed from the bank, without requiring any retraining whatsoever.

For the present study, gradient-boosting models were trained using the well-known XGBoost library [29]. Binary cross-entropy loss was minimized during both the training and tuning of each individual classifier. Training was performed using the appropriately labeled "Train 2" dataset and hyperparameter tuning was performed using the "Validation 2" dataset. The performance of each individual classifier was gauged by evaluating the model on the "Test 2" dataset and then computing the $F_1$ score [30]. The multiclass performance of the entire bank of binary classifiers was measured by selecting the class with the highest prediction probability for each sample in the "Test 1" dataset and then computing a macro-averaged multiclass $F_1$ score.

**Hyperparameter Optimization** In order to make sure the performance of each proposed band set was being maximized, the hyperparameters of each XGBoost model were tuned using a Bayesian optimization routine, implemented via the Optuna package [31].

### 2.2.5 Current Band-Set Adjustment

In addition to the training and tuning of binary-classification models, the band-optimization module (highlighted in blue in Fig. 7) relies on a spectral-downsampling procedure and an optimizer.

[6]It may be noted that nearly all recent investigations into ML-driven RSO characterization have relied heavily on DL-driven approaches. The authors of the present study have found that deep neural networks do not provide a meaningful improvement in performance over simpler ML-based approaches when attempting to classify pixelwise spectra.

**Spectral Downsampling** Since all the training spectra have been generated at hyperspectral resolution, a spectral-downsampling process is applied prior to retraining the classification models at any new set of band centers and bandwidths. This is accomplished by representing the spectral-response function of each suggested band as downward-opening parabola with its vertex corresponding to a height of unity. The suggested center wavelength is set as the vertex and the suggested bandwidth is used to set the full width at half maximum (FWHM). Multiplying each hyperspectral training sample by each of the newly suggested spectral-response functions and then summing the result constitutes a physically accurate way of interpolating the original, high-resolution spectral signatures down to an arbitrarily defined set of multispectral bands. This downsampling process was applied at each iteration of the band-optimization module.

**Constrained Optimization** At each iteration of the band-optimization module, a Bayesian-optimization algorithm was used to refine the initial guesses supplied by the importance-weighted peak-finding process. Note that it is important to use a derivative-free optimizer for such a study, as analytical gradients of the underlying cost function (i.e., the multiclass performance of the bank of binary classifiers) cannot be derived. It is well to mention that the measure of multiclass performance used as a cost function here is not the same as that used for evaluating overall performance described in Sec. 2.2.4. Instead of computing the macro-averaged $F_1$ score, the cost function seeks to minimize the multiclass cross-entropy. In order to avoid biasing the optimization results, the cost function is evaluated using the "Validation 1" dataset.

The optimization process was provided constraints representing real-world design considerations. Specifically, the following constraints were enforced:

- To avoid complications related to opacity in the ultraviolet regions of the spectrum, no center wavelengths below 200 nm were permitted
- To accommodate practical design considerations, no bandwidth was permitted to be narrower than 50 nm or wider than 200 nm
- Following a general rule of thumb for spectral-filter design, no bandwidth was permitted to exceed a quarter of the value of its corresponding center wavelength.

### 2.2.6 Novel Band-Set Proposal

As described in Sec. 2.2.2, the band-proposal module (highlighted in green in Figs. 6 and 7) uses XAI techniques to compute the importance of each band and then relies on a smoothing and peak-finding process to generate a new set of candidate bands.

**Band Importance** Although there are several XAI techniques [32] that can be used to recover feature-importance values, XGBoost models natively include a feature-importance attribute. For the present study, the "gain" mode is used to compute feature importance. This mode measures the aggregate improvement to the loss function across all splits made on any given feature.

**Importance-Spectrum Smoothing** In order to remove noise and spurious peaks from the so-called "importance spectrum" (i.e., the summed plot of importance-weighted spectral-response functions), a smoothing operation is applied. The smoothing operation, which effectively applies a Gaussian filter, comprises multiple processing steps:

1. Count the number of local maxima, $n_{max}$, in the importance spectrum and make a note of the maximum importance value, $i_{max}$
2. Zero pad the importance spectrum and take the Fourier transform
3. Compute the cutoff frequency, $f_c$, corresponding to a $\log_{10} n_{max}$ percentage attenuation of the energy contained in the alternating-current frequencies
4. Compute a Gaussian function with zero mean and $\left(\frac{f_c}{2}\right)^2$ variance

Table 2: The center wavelengths and corresponding bandwidths (FWHM values) of the five optimized multispectral bands seen in Fig. 9b are listed.

| **Band Number** | **Center Wavelength** (nm) | **Bandwidth** (nm) |
|---|---|---|
| 1 | 2,458 | 115 |
| 2 | 3,331 | 101 |
| 3 | 4,434 | 139 |
| 4 | 6,489 | 76 |
| 5 | 8,464 | 197 |

5. Multiply the Fourier transform by this Gaussian and take the inverse Fourier transform
6. Scale the result to correspond to $i_{max}$.

**Peak-Finding** Peaks in the importance spectrum and their corresponding widths were found using an algorithm based off the `peak_widths` function [33]. This algorithm is capable of returning the locations of local maxima as well as providing an estimate of the width of each maximum. If necessary, the bandwidths are clipped to adhere to the filter-design constraint described in Sec. 2.2.5. These widths are used as initial guesses for the FWHM values of each newly proposed band.

## 3. RESULTS

A notional design study was performed to (a) find the minimal set of multispectral bands required for the generic SDA mission described in Sec. 1.1 and (b) train a bank of ML models for classifying RSOs using spectral signatures defined only at that minimal set of bands.

### 3.1 Optimization Process Summarized

Supplied with the original set of 1,574-band hyperspectral signatures (see Sec. 2.1.1), the optimal band-design procedure was able to find a set of five multispectral bands within five iterations. This represents a 99.7% reduction in the number of required spectral bands. Despite such a sharp reduction in the number bands, however, the multiclass performance (measured using the macro-averaged $F_1$ score) only decreased by 1.9%, from 0.8920 to 0.8747. The change in both the macro-averaged $F_1$ score as well as the number of bands over the course of the five design iterations is presented in Fig. 8.

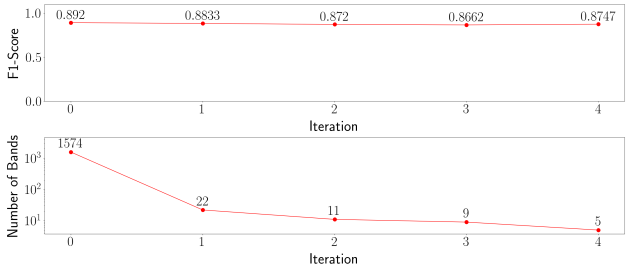
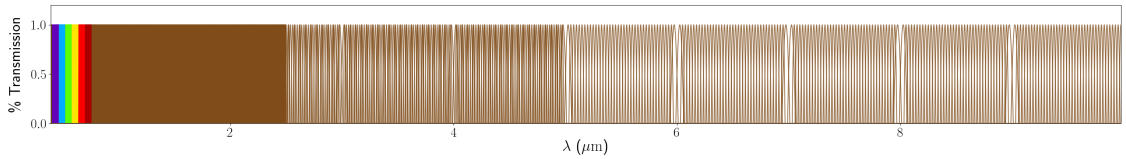



Fig. 8: When applied to a notional SDA mission, the optimal band-design procedure converged within five iterations. (The final iteration, when the stopped criterion was met, is not recorded here.) The top plot shows the evolution of the multiclass $F_1$ score over the course of those five iterations; the bottom plot shows the evolution of the number of bands. N.B. The ordinate values in the bottom plot are graphed on a logarithmic scale.

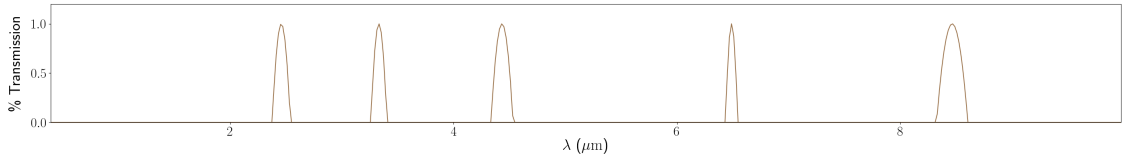
The spectral-response functions of the original 1,574 bands and the final five bands are shown in Fig. 9. In these plots, visible regions of the spectrum are colored as they would appear to the human eye, whereas IR regions are colored maroon. All five of the optimized bands found, whose parameters are listed in Table 2, fall within the IR portion of the spectrum.

### 3.2 ML Classifier Performance

As seen in Fig. 8, the overall multiclass $F_1$ score was only reduced by 1.9% over the course of the optimization process. This was due to the binary $F_1$ scores for the three classifiers remaining largely constant during the design process. In Fig. 10, the change in classifier performance between the start and end of the optimization procedure is plotted. Intuition suggests that models trained on lower-resolution spectra ought to yield lower classification performance. As seen in the figure, however, this is not always the case. When the original hyperspectral signatures were downsampled to the five bands returned by the optimal band-design procedure, the impact on classifier performance was negligible. In fact, for the active-payload class, there was no change in $F_1$ score whatsoever. The confusion matrices presented




(a) Original hyperspectral bands




(b) Optimized multispectral bands

Fig. 9: The two subfigures show the spectral-response functions associated with the original, hyperspectral band set (comprising 1,574 bands) and the optimized, multispectral band set (comprising five bands) returned by the band-design procedure.

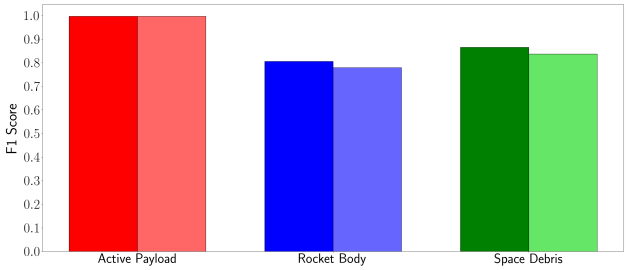



Fig. 10: The change in $F_1$ score for the ML classifiers corresponding to each RSO class over the course of the five design iterations. For each class, the darker bar on the left shows the $F_1$ score of the binary classifier trained on spectra defined at the original 1,574 bands; the lighter bar on the right shows classifier performance when trained on spectra defined at the five optimized bands.

in Fig. 11 tell a similar story. The overall, multiclass performance of the bank of trained ML classifiers is nearly unaffected between the start of the procedure and the end.

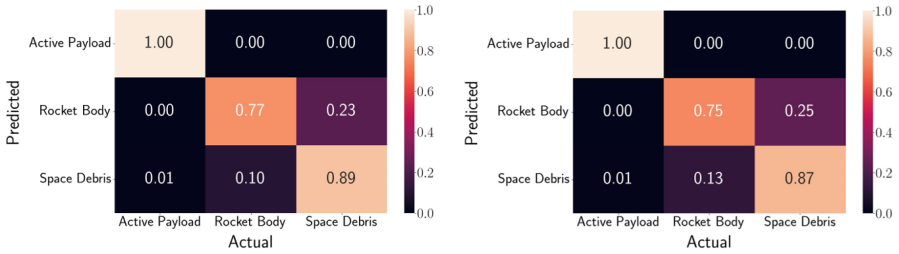
### 3.3 Iteration Details

This section presents the processing performed by the band-proposal module for each of the five iterations. The results of the process are shown as a series of figures. As recorded in Fig. 8, the number of bands was reduced monotonically: 1,574 $\rightarrow$ 22 $\rightarrow$ 11 $\rightarrow$ 9 $\rightarrow$ 5 $\rightarrow$ 5 (stopping criterion).

The process by which each of these reductions was achieved is presented in Figs. 12 through 16. Each of these figures comprises five subplots. The topmost plot shows the sum of feature-importance values found across all three ML classifiers trained and tuned during the band-optimization module. The second plot marks the normalized values of those sums explicitly for each band present. The third plot shows the spectral-response functions after having been weighted by the normalized feature-importance sums. In this plot, the notation "%T" denotes *percentage transmission*. The plot appearing second from the bottom shows the importance spectrum, i.e., the summation of the importance-weighted spectral-response functions. Lastly, the bottommost plot shows peaks being identified along the smoothed importance spectrum.

## 4. CONCLUSION

This paper has presented an ML-informed process for optimally designing mission-relevant multispectral sensors. The *optimal band-design procedure* accepts hyperspectral signatures representing classes of interest and returns (a) a minimal set of spectral bands required for detecting those classes and (b) a trained bank of ML models capable of classifying spectra sampled at that minimal set of bands. A design study was performed for a generic SDA mission,




(a) Trained on 1,574-band spectra (b) Trained on 5-band spectra

Fig. 11: The two subfigures show the confusion matrices for the full bank of ML classifiers, computed at the beginning and at the end of the band-design procedure.

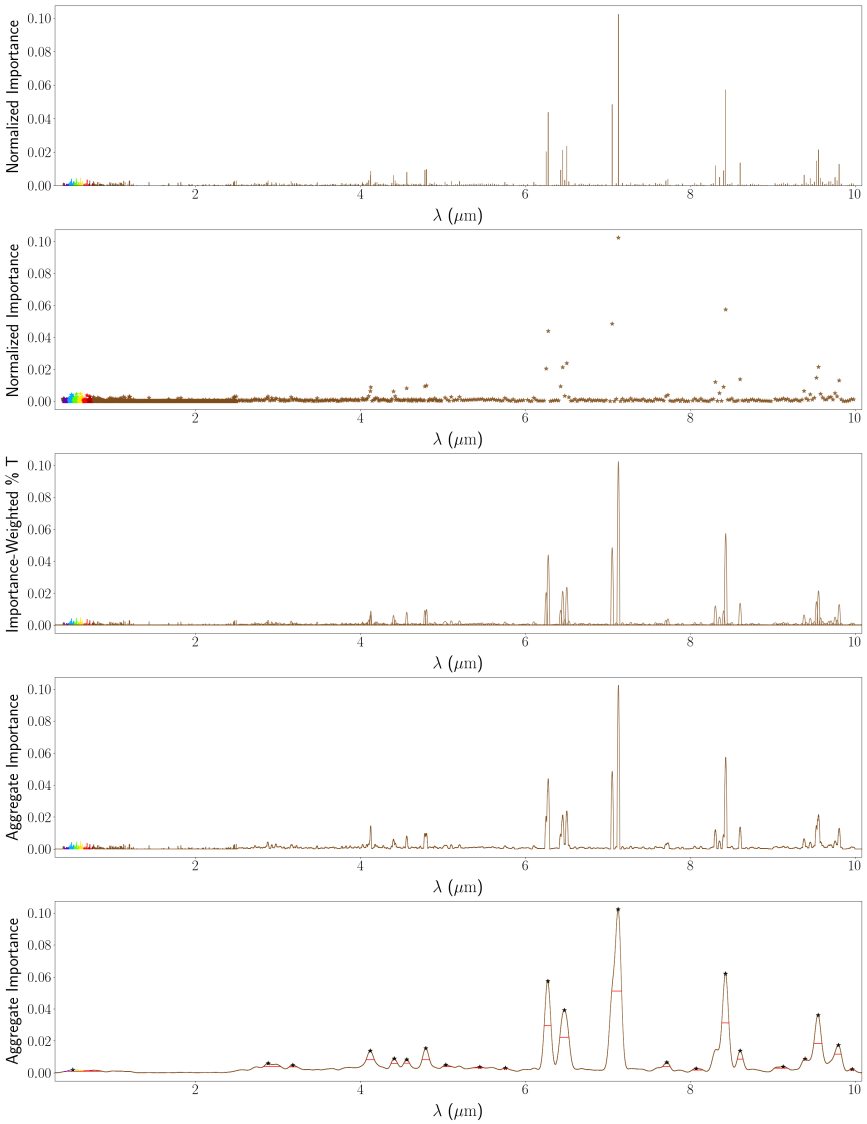



Fig. 12: Iteration 1: 1,574 bands → 22 bands. The largest per-iteration reduction in bands (98.6%) was seen at the very start of the procedure. The effects of the smoothing operation are especially noticeable, where several small local maxima are suppressed.

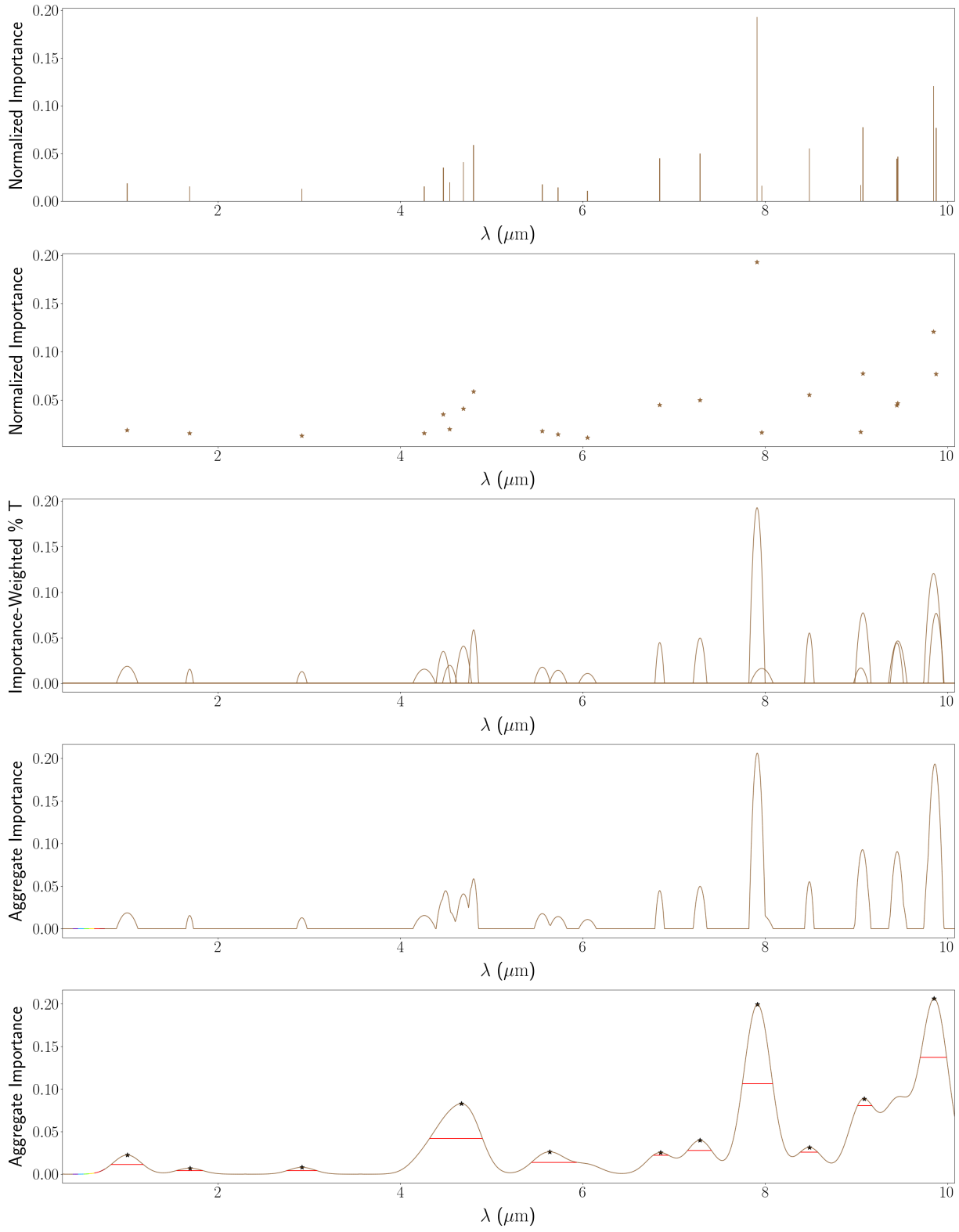



Fig. 13: Iteration 2: 22 bands → 11 bands. The second iteration saw a 50% reduction in the number of bands. Despite the fact that there are overlapping bands at the start of this iteration, clear peaks in the importance spectrum are beginning to emerge by the end.

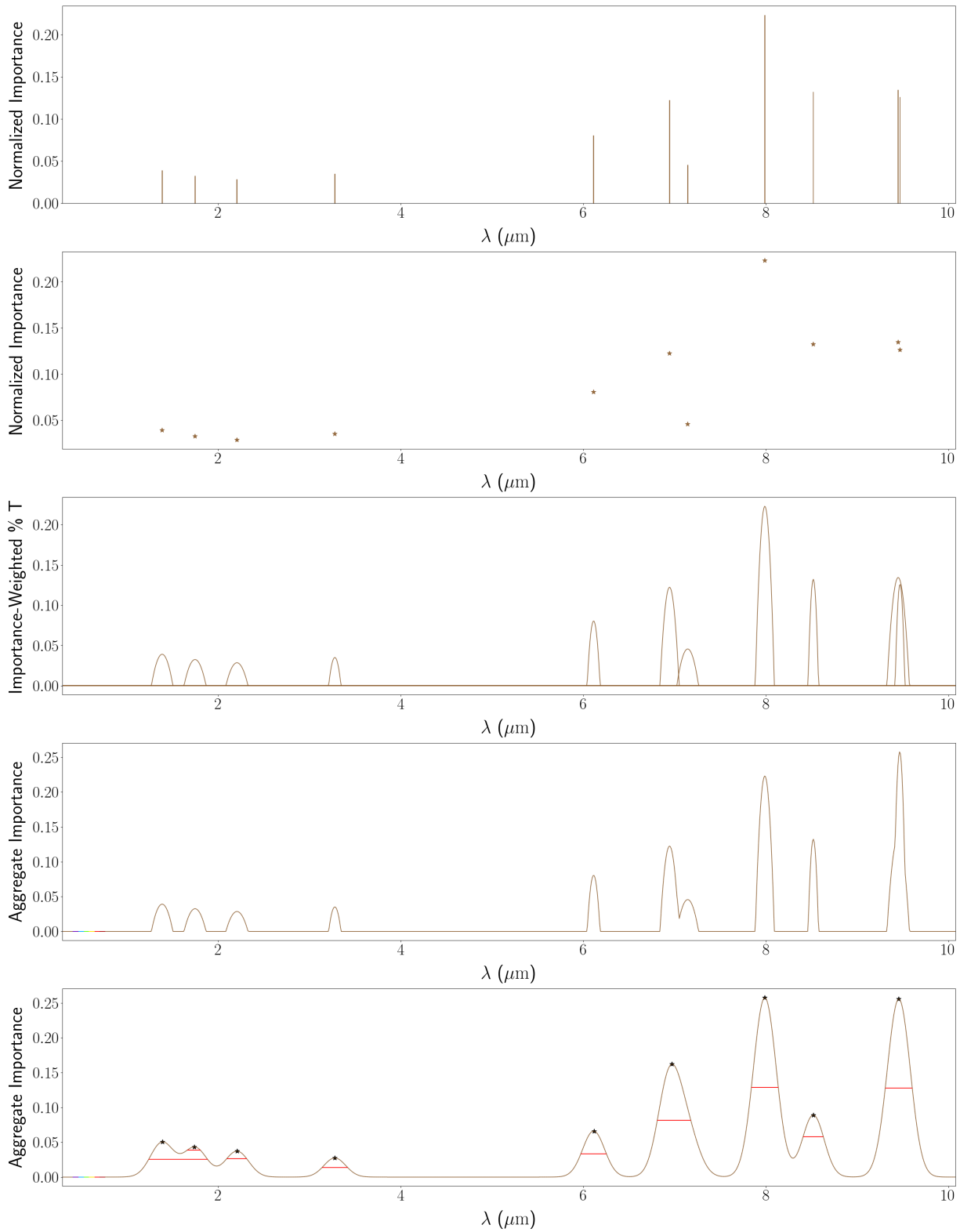



Fig. 14: Iteration 3: 11 bands → 9 bands. The third iteration is responsible for an 18.2% reduction in the number of bands. N.B. All overlapping bands have been eliminated by the end of this iteration.

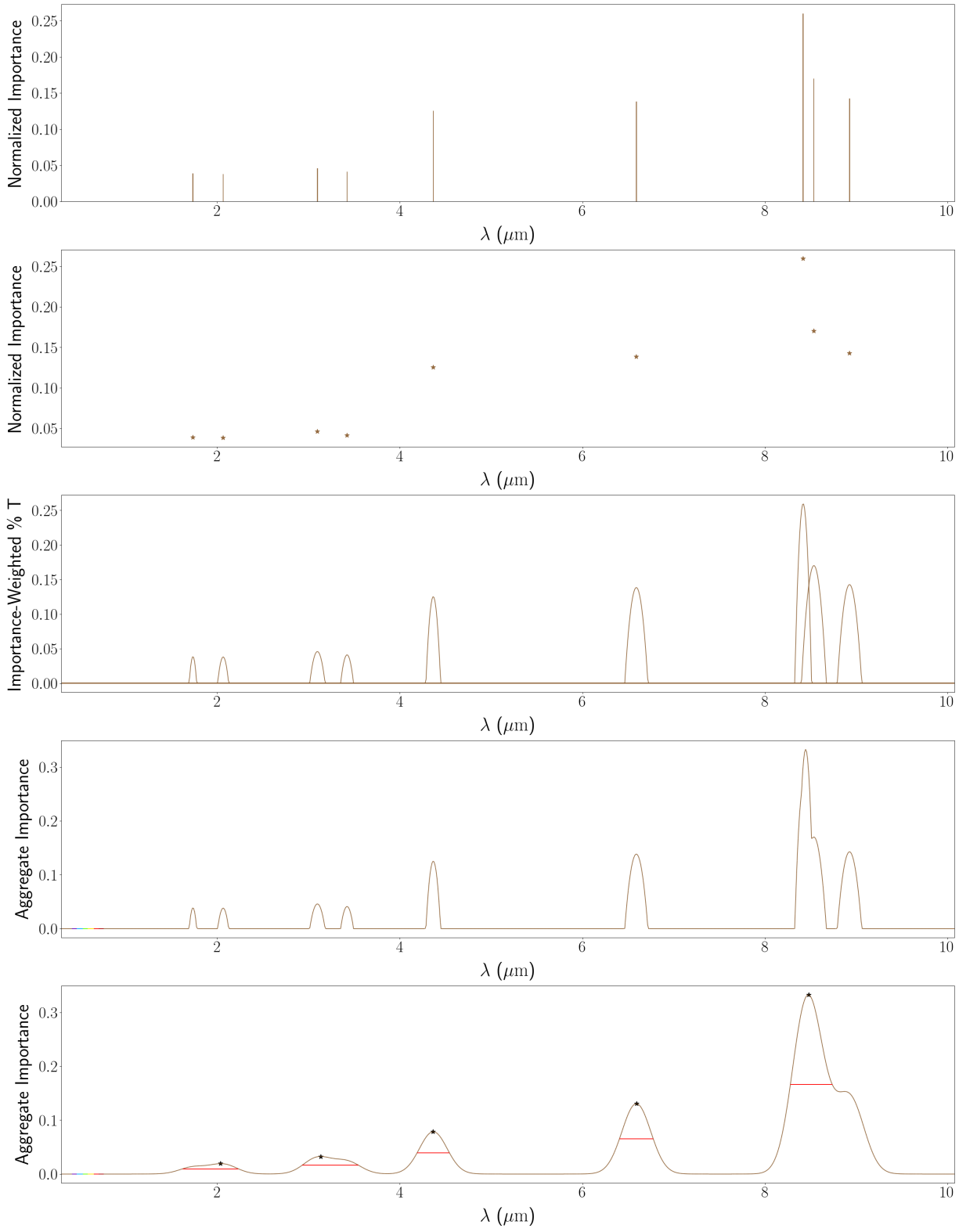



Fig. 15: Iteration 4: 9 bands → 5 bands. The fourth iteration once again results in a sharp reduction in the number of bands (44.4%). At the end of this iteration, five distinct and well-separated peaks have appeared in the smoothed importance spectrum.

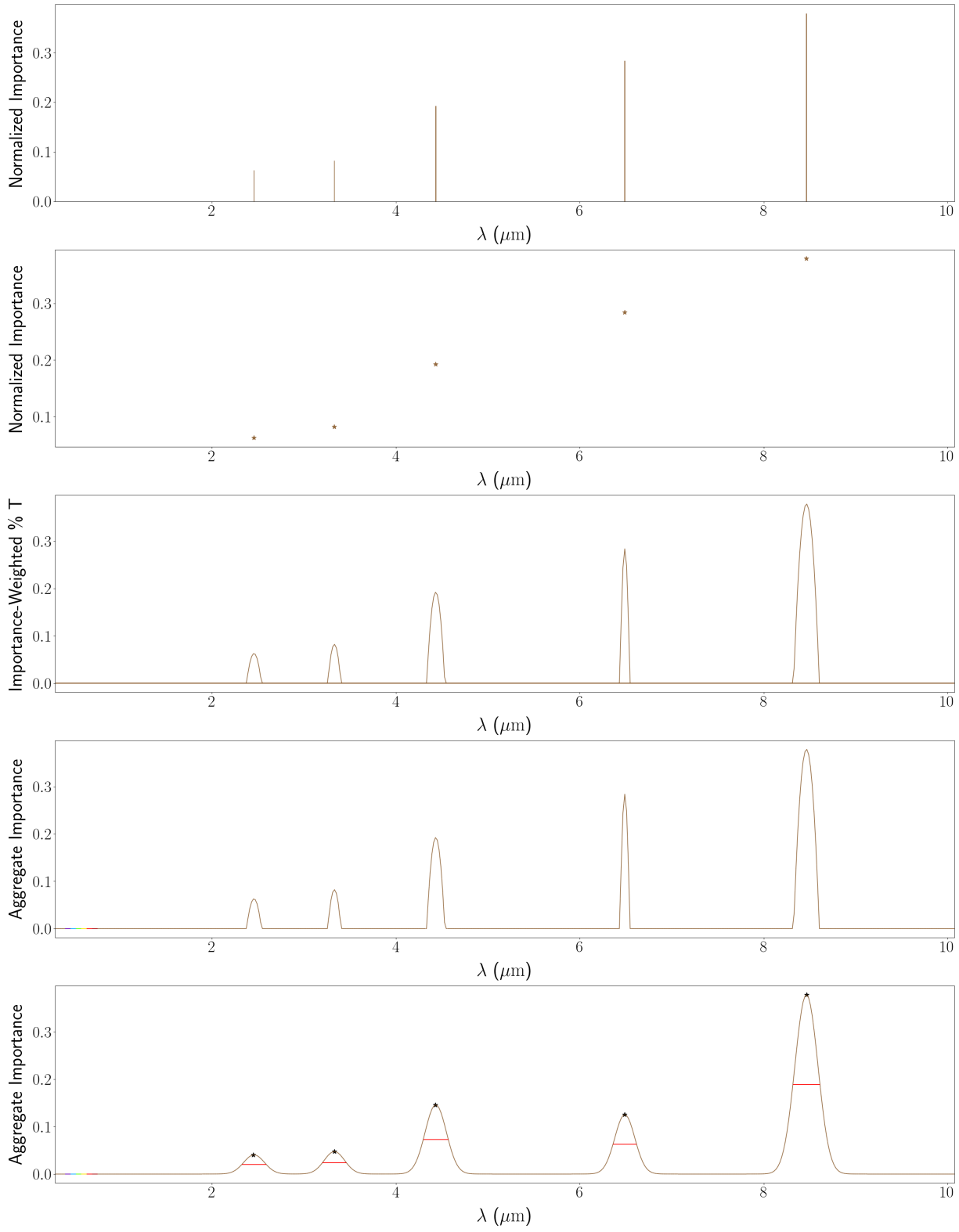



Fig. 16: Iteration 5: 5 bands → 5 bands. As the fifth iteration is unable to further reduce the number of bands, the stopping condition is triggered. The five bands found at the end of the previous iteration are clipped (see Sec. 2.2.6) and then returned.

seeking to identify three classes of RSO. Starting from hyperspectral signatures defined at 1,574 bands, the procedure was able to design a multispectral system with only five bands. This 99.7% decrease in the number of bands came at the cost of only a 1.9% decrease in the performance of the ML classifiers. This study suggests that future SDA missions could realize SWaP-C savings by replacing hyperspectral sensors with bespoke multispectral sensors without sacrificing overall performance.

## 5. ACKNOWLEDGEMENTS

The authors would like to thank the SDA mission architects and spectral analysts at BAE Space & Mission Systems who assisted in the preparation of this paper, namely: Geoff Lake, Naomi Owens-Fahrner, Anna Lawitzke, Bill Baugh, and Matt Fisher. They would also like to thank the managerial and legal staff at BAE Space & Mission Systems, including Janet Phillips, Denise Waller, and Cynthia Wallace, all of whom deserve recognition here for their assistance with procedural details and security compliance.

## 6. REFERENCES

[1] M. J. Holzinger and M. K. Jah. Challenges and potential in space domain awareness. *Journal of Guidance, Control, and Dynamics*, 41(1):15–18, 2018.

[2] Konstantinos Tsaprailis, George Choumos, Vaios Lappas, and Charalampos Kontoes. Survey mode: A review of machine learning in resident space object detection and characterization. In *AIAA SCITECH 2024 Forum*. AIAA, 2024.

[3] The European Space Agency. ESA — About space debris. `https://www.esa.int/Space_Safety/Space_Debris/About_space_debris`, 2024. (Accessed: 10 August 2024).

[4] Michael Byers and Aaron Boley. *Abandoned Rocket Bodies*, page 114–129. Cambridge Studies in International and Comparative Law. Cambridge University Press, 2023.

[5] Michael J. Nicolls and Darren Mcknight. Collision risk assessment for derelict objects in low-earth orbit. In *First International Orbital Debris Conference*, 2019.

[6] Heather Cowardin, Pat Seitzer, Kira Abercromby, Ed Barker, and Thomas Schildknecht. Characterization of orbital debris photometric properties derived from laboratory-based measurements. In *AMOS Technical Conference*, number JSC-CN-21447 in 2010, 2010.

[7] Ryan Clark, Yanchun Fu, Siddharth Dave, and Regina S K Lee. Resident space object (rso) attitude and optical property estimation from space-based light curves. *Advances in Space Research*, 70(11):3271–3280, 2022.

[8] Rachel O'Connor, E. Natasha Stavros, Madeline Cowell, Nathan Leisso, Lyle Ruppert, Betsy Maria Farris, Nathan Showalter, Carl S. Weimer, and William Baugh. Performance Comparisons of Prism-Based and Grating-Based Airborne and Spaceborne Hyperspectral Spectrometers. In *AGU Fall Meeting Abstracts*, pages GC51G–0696, dec 2023.

[9] Timothy Stratman and Shawana P. Johnson. Novel 2d hyperspectral imaging, delivering a new realm of information for a variety of next generation applications. In *ASCEND 2020*. AIAA, 2020.

[10] Kira Abercromby, Jennifer Okada, Michael Guyote, Kris Hamada, and Edwin Barker. Comparisons of ground truth and remote spectral measurements of the formosat and ande spacecraft. In *Proceedings of the 2006 AMOS Technical Conference*. Citeseer, 2006.

[11] Jacqueline Reyes and Darren Cone. Characterization of Spacecraft Materials using Reflectance Spectroscopy. In S. Ryan, editor, *The Advanced Maui Optical and Space Surveillance Technologies Conference*, page 57, September 2018.

[12] Kira Jorgensen, John L Africano, Eugene G Stansbery, Paul W Kervin, Kris M Hamada, and Paul F Sydney. Determining the material type of man-made orbiting objects using low-resolution reflectance spectroscopy. In *Multifrequency Electronic/Photonic Devices and Systems for Dual-Use Applications*, volume 4490, pages 237–244. SPIE, 2001.

[13] Alessandro Vananti, Thomas Schildknecht, Holger Krag, and Christian Erd. Preliminary results from reflectance spectroscopy observations of space debris in geo. In *Proc. 5th European Conf. on Space Debris. ESA SP-672. ESA/ESOC. Darmstadt, Germany*, 2009.

[14] Alessandro Vananti, Thomas Schildknecht, and Holger Krag. Reflectance spectroscopy characterization of space debris. *Advances in space research*, 59(10):2488–2500, 2017.

[15] Jainisha Shah, Miles T. Bengtson, Elena Plis, Ryan C. Hoffmann, Dale C. Ferguson, Sydney Collman, Daniel P. Engelhart, Gregory Badura, Heather Cowardin, and Timothy Scott. Spacecraft material characterization using reflectane spectra extracted from rgb/ir color images. In *AIAA SCITECH 2023 Forum*. AIAA, 2023.

[16] Matthew R. Radcliff. NASA SVS — Lansdsat sensors: pushbroom vs whisbroom. `https://svs.gsfc.nasa.gov/12754/`, 2017. (Accessed: 24 August 2024).

[17] Aryzbe Najera, Miguel Velez-Reyes, Hector Erives-Contreras, and Dan DeBlasio. Simulating ground-based measurements of spectral signatures of resident space objects. In Miguel Velez-Reyes and David W. Messinger, editors, *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXIX*, volume 12519, page 1251908. International Society for Optics and Photonics, SPIE, 2023.

[18] Adam A. Goodenough and Scott D. Brown. DIRSIG5: Next-generation remote sensing data and image simulation framework. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11):4818–4833, 2017.

[19] F. Hernandez, S. Almond, and M. Li. Hyperspectral Imaging Analysis of Simulated Scenes for Space Domain Awareness. In S. Ryan, editor, *Proceedings of the Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference*, page 60, September 2023.

[20] Eric Coiro, Ugo Tricoli, François Margall, and Cyril Petit. Hyperspectral optical modeling of resident space objects at high spatial resolution. In Miguel Velez-Reyes and David W. Messinger, editors, *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXX*, volume 13031 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 130310A, June 2024.

[21] Xiao-Fen Zhao, Hui-Yan Zhang, Yong Yu, and Yin-Dun Mao. Multicolor photometry of geosynchronous satellites and application on feature recognition. *Advances in Space Research*, 58(11):2269–2279, 2016.

[22] Saziye Ozge Atik and Muhammed Enes Atik. Optimal band selection using explainable artificial intelligence for machine learning-based hyperspectral image classification. *Journal of Applied Remote Sensing*, 18(4):042604, 2024.

[23] Kedar R. Naik, Andrew I. Wernersbach, Michelle F. Nilson, Matthew D. Fisher, William M. Baugh, and Gary D. Wiemokly. Hyperspectral to multispectral: optimal selection of mission-relevant bands using machine learning. In Miguel Velez-Reyes and David W. Messinger, editors, *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXX*, volume 13031, page 1303107. International Society for Optics and Photonics, SPIE, 2024.

[24] Matthew Phelps, J. Zachary Gazak, Thomas Swindle, Justin Fletcher, and Ian Mcquaid. Inferring space object orientation with spectroscopy and convolutional networks. *AMOS (Sept. 2021)*, 3, 2021.

[25] J. Zachary Gazak, Ian McQuaid, Ryan Swindle, Matthew Phelps, and Justin Fletcher. Spectranet: Learned recognition of artificial satellites from high contrast spectroscopic imagery. In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 2403–2411, 2022.

[26] Na Li, Chengeng Gong, Huijie Zhao, and Yun Ma. Space target material identification based on graph convolutional neural network. *Remote Sensing*, 15(7), 2023.

[27] Massimiliano Vasile, Lewis Walker, Andrew Campbell, Simão Marto, Paul Murray, Stephen Marshall, and Vasili Savitski. Space object identification and classification from hyperspectral material analysis. *Scientific Reports*, 14, 01 2024.

[28] Luis R. Cedillo, Kevin Acosta, Miguel Velez-Reyes, and Dan DeBlasio. Using neural networks to classify hyperspectral signatures of unresolved resident space objects. In Miguel Velez-Reyes and David W. Messinger, editors, *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXX*, volume 13031 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 130310B, June 2024.

[29] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[30] C. J. Van Rijsbergen. *Information Retrieval*. Butterworth-Heinemann, USA, 2nd edition, 1979.

[31] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631, 2019.

[32] Wenzhuo Yang, Hung Le, Silvio Savarese, and Steven Hoi. OmniXAI: A Library for Explainable AI, 2022.

[33] The SciPy Community. scipy.signal.peak_widths – SciPy v1.13.0 Manual. SciPy `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.peak_widths.html`, 2023. (Accessed: 03 April 2024).