SF2A 2025
A. Siebert, K. Baillié, M. Béthermin, F. Cantalloube, E. Josselin, N. Lagarde, J. Malzac, J. Richard, L. Selliez, O. Venot (eds)

MACHINE LEARNING-BASED IDENTIFICATION AND CLASSIFICATION OF BAD PIXELS FOR THE ARIEL INFRARED AIRS DETECTOR

A. Syty[1], O. Faucoz[2], C. Cossou[3], T. Pichon[3], P. Drossart[1] and J. P. Beaulieu[1]

Abstract. The performance of space-based infrared detectors is often limited by non-nominal pixel behaviors, which can bias scientific observations if not properly identified and corrected. In this work, we present a machine learning framework to characterize and classify bad pixels in the Ariel AIRS-CH1 detector using dark-frame measurements. Unsupervised clustering with a Gaussian Mixture Model is first applied to construct a verified benchmark of pixel behaviors, which is then used to train a Random Forest algorithm for full-detector classification. The resulting pixel maps reveal spatial patterns of defective pixels, including cosmic-ray hits and non-linear responses, and provide a confidence metric for each prediction. This approach enables targeted pre-flight detector region selection and post-flight correction of well-characterized bad pixels, ensuring high-quality photometric data for exoplanet atmosphere studies.

Keywords: Ariel mission, infrared detector, bad pixel classification, machine learning, Random Forest

1 Introduction

The Ariel space mission is a European Space Agency (ESA) mission that aims to study the atmospheres of a large and diverse catalog of transiting exoplanets (Tinetti et al. 2022). Scheduled for launch in 2029 to the L2 Lagrange point, Ariel will observe a diverse catalog of transiting exoplanets in visible and near-infrared wavelengths (0.5-7.8 μm) via low-resolution spectroscopy. By analyzing starlight filtered through exoplanet atmospheres during transits, the mission will detect wavelength-dependent variations in a planet's apparent radius caused by molecular absorption.

As machine learning methods are increasingly used in many astrophysics fields, these methods have recently arrived in the exoplanet community. The Ariel Consortium puts effort into the machine learning development from data processing to molecular retrievals (Changeat & Yip 2023). Applications include accelerating parameter space exploration, investigating the highly degenerate molecular composition of planet atmospheres in exoplanets' spectra, studying complex undesired artifacts in Ariel data, such as the effect of the jitter of the line of sight of the telescope during the observation, or even interpreting pre-launch calibration data. Indeed, post-processing methods will be used to correct Ariel data up to photometric noise and ensure that the science objectives can be achieved (Bocchieri et al. 2025). These corrections often rely on the knowledge of calibration maps of the detector, acquired before flight. However, the detector's performance may vary in flight due to extreme conditions of temperature and pressure or because of strong mechanical constraints experienced during launch. The performances may also vary over the years the telescope is in operation. For instance, the ability to detect bad pixels in flight might be a key element of the success of the mission. Indeed, if left uncorrected, bad pixels can introduce bias into the data, potentially limiting the extraction of atmospheric features from exoplanet spectra. If a bad pixel is spotted, it can be either masked from the image for its analysis or corrected if its behavior is well characterized.

Dark frames are measured in AIM/CEA (Astrophysique, Instumentation, Modélisation) (Martignac et al. 2024) on the engineering model detector of AIRS-CH1 to flag the pixels having non-nominal behaviors ("bad pixels"), which must be either masked during the data processing step or parametrized to avoid introducing

[1] Institut D'astrophysique de Paris (IAP), Sorbonne University, CNRS, 75014 Paris, France
[2] Centre National d'Études Spatiales (CNES), 31400 Toulouse, France
[3] Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), Paris-Saclay University, 91191 Gif-sur-Yvette, France

any bias in the scientific data to come. Machine learning is a powerful tool to cluster those pixels according to their behavior, on datacubes containing millions of time series of pixels' responses. Different data pre-processing methods, as well as clustering methods such as Gaussian Mixtures and Random Forests, are used to classify and characterize the various types of bad pixels.

2 Methods

2.1 Dark-Frame Measurements

Dark frames are measurements of the detector's response without being illuminated by any external source. The signal acquired is the noise from the electronics, accumulating over time. We work on dark frames measured of two different HgCdTe H1RG detectors from Teledyne Imaging System: a infrared test detector and the Engineering Model (EM detector) of Ariel AIRS-CH1 infrared spectrometer.

The objective of the method is both to discriminate bad pixels from good ones and to cluster bad pixels according to their type of behavior. This characterization of pixel responses is necessary to parametrize or correct pixels that are non-nominal but still contain usable scientific information. A good pixel is defined as one that integrates charges linearly over time. Various types of bad pixels can be found in HgCdTe detectors. Dead pixels show no response over time, while hot pixels exhibit abnormally high dark current, producing a strong signal even without illumination, typically with stable behavior. Some pixels display Random Telegraphic Noise, characterized by discrete and random switching between two or more signal levels, often associated with charge trapping and detrapping processes in the detector material. RC pixels show a highly non-linear response, and saturated pixels reach the maximum charge capacity of the well very quickly, remaining at a constant high output level regardless of further illumination. Pixels affected by Cosmic Rays exhibit sudden jumps in their signal ramp due to excess charge deposited in the detector material. Other pixels may show unexpected behaviors, such as decreasing signals over time.

2.2 Reducing the parameters' space

The first step of the clustering method is to find a set of parameters that contains most of the information but with a lower dimensional complexity. Several technics can be used to reduce the size of the parameters' space, such as principal components analysis (PCA) or encoding the information in a reduced latent space using autoencoders. However, we wanted the choice of parameters used to be motivated by physical arguments. Therefore, we decided to compute for each pixel:

- The mean, standard deviation, skewness and kurtosis of the time distribution of the signal.
- The sum of variations, mean, standard deviation, median, maximum absolute value, the ratio of the maximum to the mean of the others, the number of jumps above an adaptive threshold, skewness, and kurtosis of the rate of change of the signal. The rate of change formula is the following :

$$ROC = \frac{s_{n+1} - s_n}{t_{n+1} - t_n}$$

  s being the signal value and t the various time steps of the measurement.
- The five first coefficients of a wavelet transform, bringing more time-correlated information in the parametrization. We chose to keep only the five first coefficients as they contained most of the information needed to describe the signal over time.

The different extracted parameters are then normalized using a global min-max normalization to ensure they are all between 0 and 1.

2.3 Unsupervised machine learning to create a benchmark

In the reduced parameter space, we applied a Gaussian Mixture Model (GMM) to identify clusters of points with similar characteristics. The GMM estimates the mean and standard deviation of Gaussian distributions and assigns to each data point a probability of belonging to these distributions, thereby grouping points according to their behavioral similarity within a probabilistic framework.

Other unsupervised clustering techniques were also tested, such as density-based methods like DBSCAN, but they proved ineffective due to the highly continuous distribution of points in the parameter space. In particular, some defective pixels exhibit intermediate behaviors between categories, which leads density-based algorithms to either merge all points into a single cluster or classify most of them as outliers.

The GMM provided an initial effective clustering of defective pixels. From each cluster, we randomly selected several hundred pixels to construct a benchmark dataset of bad pixels, to which we added two-thirds of good pixels. Each pixel in this dataset was manually verified to ensure the accuracy of the ground truth. This benchmark then enabled further analyses on other observations using a pre-trained Random Forest algorithm.

2.4 Supervised machine learning with Random Forests

With the benchmark dataset labeled and validated, a Random Forest model can be trained. The objective of this second step is to obtain more reliable predictions and to quantify the uncertainty of the method. The Random Forest algorithm operates by aggregating a large number of predictions from an ensemble of decision trees. Each tree is trained on a random subsample of the labeled dataset and considers a distinct subset of the parameters defined in 2.2. This training process is illustrated in the left side of Figure 1.

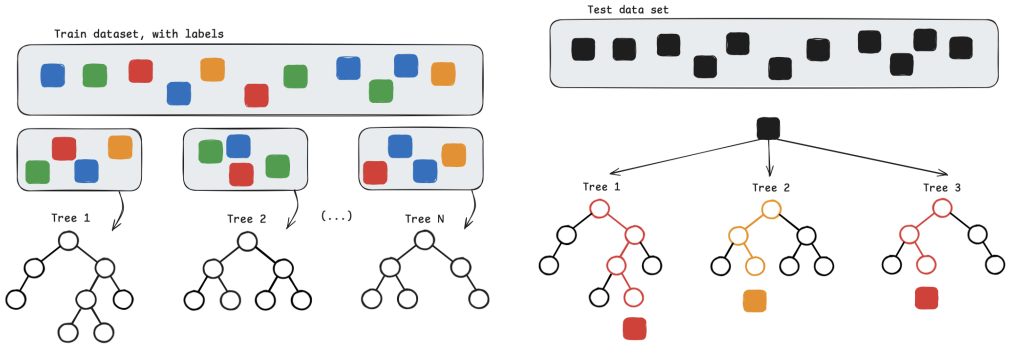

Fig. 1. Illustration of the process of training a Random Forest algorithm. On the left, the colored inputs represent pixels from the benchmark dataset with known ground truth -behavioral type. On the right, the black inputs represent pixels from a new measurement from which we don't know any ground truth. Each tree makes a prediction, here red or orange and the majority label (red) is the final prediction with a confidence level of 2/3.

Once the model has been trained, it is saved and can be applied to new datasets that have undergone the same pre-processing steps as the training data to make predictions. Each decision tree in the ensemble provides a vote, and the cluster receiving the majority of votes is assigned as the final prediction. The proportion of trees supporting the majority prediction provides a measure of the method's uncertainty. This process is illustrated on the right side of Figure 1.

3 Results

The Random Forest model predicts the cluster assignment of each pixel in the detector based on their behavior in the dark-frame measurement, along with the associated prediction confidence, as described in Section 2.4.

The resulting detector map, where each label is displayed in a different color, provides insight into the spatial distribution of the detector's characteristics. This is illustrated in Figure 2.

Figure 2 shows that some bad-pixel behaviors form spatial patterns, such as the green pixels clustering into ball-like shapes. As illustrated in Figure 3, these green pixels are affected by Cosmic Rays, which typically impact groups of neighboring pixels on the detector.

Non-linear pixels-colored purple, black, or blue depending on the degree of non-linearity-also tend to exhibit spatial clustering. In the case of non-linear pixels, the effect is often associated with higher signal levels compared to good pixels: nearly saturated pixels may cause leakage and charge trapping, which can impact adjacent pixels with crosstalk. This suggests, more generally, that when one pixel is defective, neighboring pixels should be monitored carefully, as they may also be affected.

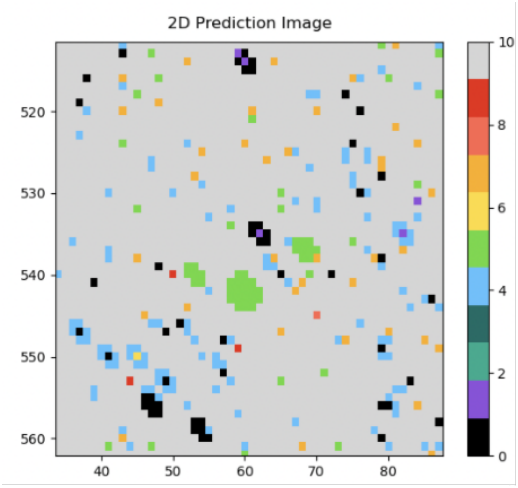

Fig. 2. This image shows a zoomed-in portion of the test detector, where pixels are colored according to their assigned cluster. Pixels in gray correspond to those with a linear evolution over time, while the other colors represent different types of identified behaviors.

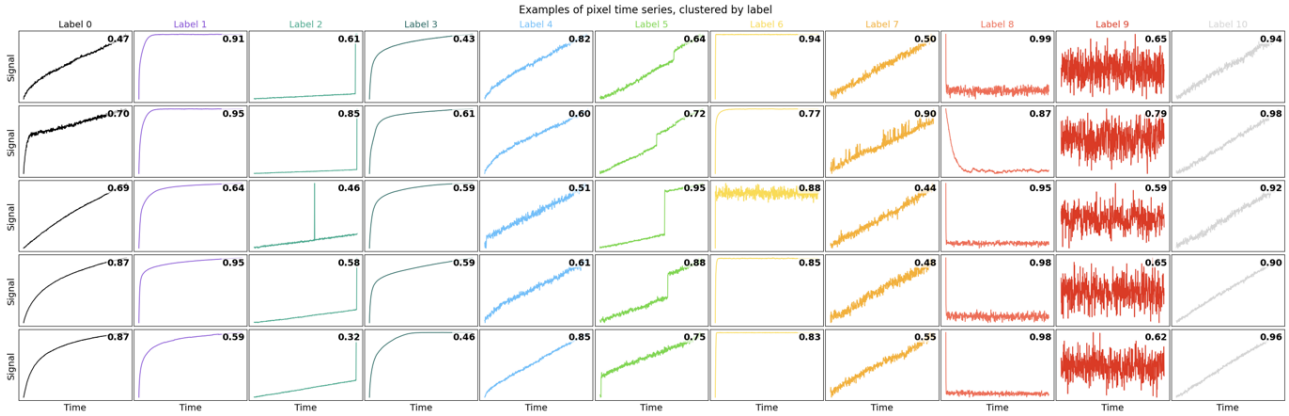

Fig. 3. Examples of ramps of pixels' signals over time, from the test detector. Each column corresponds to a different label, thus a different type of behavior. The colors are matching the one displayed on Figure 2. The confidence level of the prediction is indicated for each pixel in the upper right corner of the plot. All the signals are normalized to the same scale for better visualization.

The prediction confidence provides an indication of the reliability of the model, but more importantly, it can be used to flag pixels for which the prediction is uncertain. These pixels may exhibit uncommon behaviors that do not fit neatly into any defined category.

For this method, the overall confidence level is displayed on Figure 4.

The confidence level exceeds 90% for most pixels. However, when examining the confidence within individual clusters, some clusters show lower values. For example, as illustrated in Figure 3 for clusters 0, 1, and 3, the non-linear pixels exhibit behaviors that are not clearly separable. Even by visual inspection, it can be difficult to determine whether a pixel belongs to the highly non-linear, slightly non-linear, or intermediate cluster. This continuity in the behavior of certain bad pixels likely explains the lower confidence levels observed for clusters that are more ambiguous.

4 Conclusions

In this work, we have developed a machine learning framework to identify and classify bad pixels in the Ariel AIRS-CH1 infrared detector using dark-frame measurements. By combining unsupervised clustering with a Gaussian Mixture Model and a supervised Random Forest algorithm, we were able to construct a benchmark

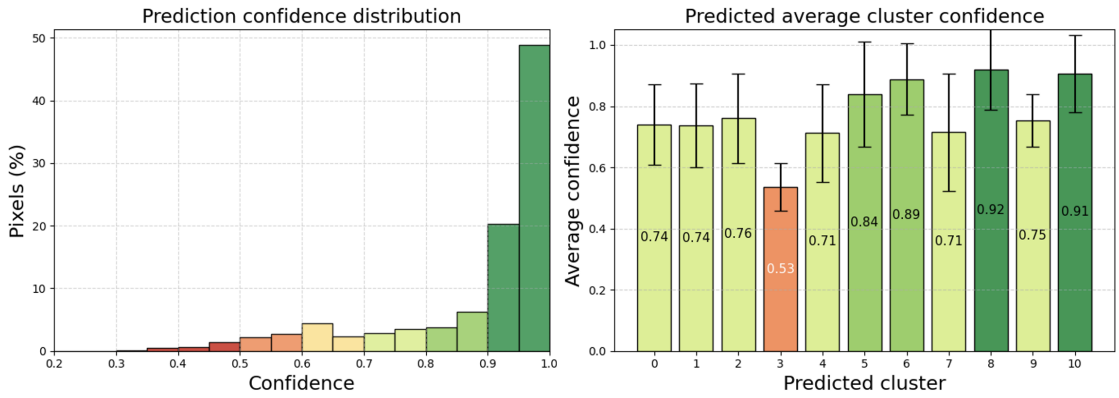

Fig. 4. Confidence level provided by the Random Forest algorithm on the test detector. On the left, the confidence levels are displayed for all pixels, and on the right, according to each cluster of interest.

dataset of verified pixel behaviors and extend the classification to other detector's measurement with a pre-trained model.

The resulting pixel maps reveal not only the spatial distribution of different types of bad pixels, but also subtle patterns, such as spatial correlations among non-linear or nearly saturated pixels. The Random Forest model provides both a predicted label and an associated confidence level for each pixel, which serves as a quantitative indicator of reliability and allows identification of pixels with ambiguous or uncommon behaviors.

This approach enables robust correction or masking of defective pixels, which is critical for ensuring the high-quality photometric data required to achieve Ariel's scientific objectives. More generally, our methodology demonstrates the potential of machine learning for detector characterization and quality assessment in space-based instrumentation. This enhanced map of bad pixels can be used pre-flight to select the region of the detector that will receive the stellar flux, or post-flight during data processing, to correct the signals of well-characterized bad pixels.

The study described here is synthesized in a python package devoted to calibration studies of the Ariel detectors and will provide parameters files to be included in Ariel simulators and pipeline software (Mugnai et al. 2025).

This work has received support from France 2030 through the project named Académie Spatiale d'Île-de-France managed by the National Research Agency under bearing the reference ANR-23-CMAS-0041, and from the Centre National d'Études Spatiales (CNES).

References

Bocchieri, A., Mugnai, L. V., Pascale, E., et al. 2025, Experimental Astronomy, 59, 31

Changeat, Q. & Yip, K. H. 2023, RAS Techniques and Instruments, 2, 45

Martignac, J., Amiaux, J., Capocci, T., et al. 2024, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 13092, Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave, ed. L. E. Coyle, S. Matsuura, & M. D. Perrin, 130921D

Mugnai, L. V., Bocchieri, A., Pascale, E., Lorenzani, A., & Papageorgiou, A. 2025, Experimental Astronomy, 59, 9

Tinetti, G., Eccleston, P., Lueftinger, T., et al. 2022, in European Planetary Science Congress, EPSC2022–1114