A GAN approach for Anomaly Detection in Spacecraft Telemetries

Carlo Ciancarelli, Giorgio De Magistris, Salvatore Cognetta, Daniele Appetito, Christian Napoli and Daniele Nardi

Abstract In spacecraft health management a large number of time series is acquired and used for on-board units surveillance and for historical data analysis. The early detection of abnormal behaviors in telemetry data can prevent failures in the spacecraft equipment. In this paper we present an advanced monitoring system that was carried out in partnership with Thales Alenia Space Italia S.p.A, a leading industry in the field of spacecraft manufacturing. In particular, we developed an anomaly detection algorithm based on Generative Adversarial Networks, that thanks to their ability to model arbitrary distributions in high dimensional spaces, allow to capture complex anomalies avoiding the burden of hand crafted feature extraction. We applied this method to detect anomalies in telemetry data collected from a simulator of a Low Earth Orbit satellite. One of the strengths of the proposed approach is that it does not require any previous knowledge on the signal. This is particular useful in the context of anomaly detection where we do not have a model of the anomaly. Hence the only assumption we made is that an anomaly is a pattern that lives in a lower probability region of the data space.

Carlo Ciancarelli
Thales Alenia Space Italia S.p.A, Via Saccomuro, 24, Rome, 00131, Italy, e-mail: `carlo.ciancarelli@thalesaleniaspace.com`

Giorgio De Magistris, Christian Napoli and Daniele Nardi
Department of Computer, Automation and Management Engineering, Sapienza University of Rome, Via Ariosto, 25, Rome, 00185, Italy e-mail: `{demagistris,cnapoli,nardi}@diag.uniroma1.it`

Salvatore Cognetta, Daniele Appetito
Sapienza University of Rome, Via Ariosto, 25, Rome, 00185, Italy e-mail: `{cognetta.1874383,appetito.1916560}@studenti.uniroma1.it`

1 Introduction

Satellites in orbit are monitored by a network of sensors which produce a huge stream of telemetry data. When the amount of data is huge and it needs to be processed in very short time, a solution to extract meaningful information (that in our case consists in anomalous patterns) must be based on fully automatized processes. Anomaly detection in time series data is a well studied problem in both the data mining and machine learning communities. Since it always happens that normal data samples outnumber the anomalous ones, anomaly detection is considered a semi-supervised (where anomalous data is used only for testing) or unsupervised (with no information about normal or anomalous data) problem. An exhaustive review of the most common approaches can be found in [1]. In particular a general approach consists in transforming the sequences into a feature space and then use a point anomaly detection technique in the new space to detect anomalies. However this approach depends both on the anomaly detection technique and the properties of the feature space. For example clustering based methods [10][13][19] require that the anomalies do not aggregate into clusters; nearest neighbour and density based methods [6] require that the anomalies do not form dense regions in the feature space; spectral methods [7][3] assume that a projection into a different space exists such that normal and anomalous points can be clearly distinguished. Another approach consists in training a model to predict the signal in the future and then compare the predicted and observed signals to detect anomalies, like in [14]. In this paper we propose the application of the Generative Adversarial Networks (GANs) for the anomaly detection in spacecraft telemetry data. The GANs provide a deep latent representation of data that can be used directly for the assessment. In particular they implicitly extract meaningful features that can be exploited to discriminate between normal and abnormal samples through the assignment of an anomaly score. The rest of the paper is structured as follows: Section 2 briefly describes the GANs framework focusing on its application on anomaly detection; Section 3 reviews the approaches used so far to tackle the problem of anomaly detection in spacecraft telemetries; Section 4 describes our dataset while Section 5 illustrates our method along with all the implementation details; Section 6 discusses the results and the conclusion is drawn in Section 7.

2 GANs and Anomaly Detection

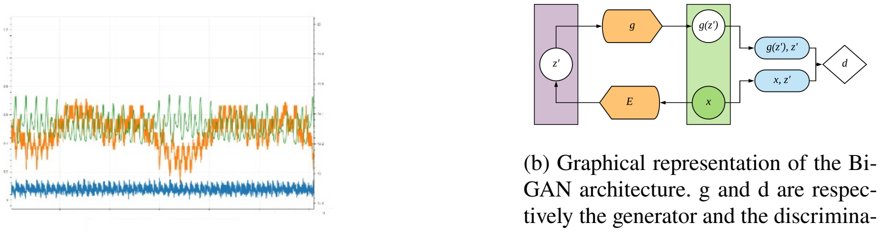
The GANs Framework, firstly introduced in [5], is composed of two networks: a generator and a discriminator, often referred as G and D. The generator learns a mapping from the latent space Θ_z, usually the set of k-dimensional standard normal vectors, to the data space Θ_{data} and the discriminator learns a categorical probability distribution over the generated and real samples to discriminate between real and fake samples. G and D optimize the same criterion in opposite directions, following a two player minimax game. The general idea behind the application of GANs to anomaly detection consists in using the output of the discriminator and the reconstruction error to assign an anomaly score to a data sample x. The output of the discriminator is the probability that a sample is “real", hence its inverse can be directly interpreted as an anomaly score. The second part is less obvious and its justification relies on the structure of the latent space. It was shown [17] that the space learned by the generator has smooth transitions, because walking on the learned manifold results in semantic changes to the generated image. This encouraged the usage of the GANs framework as an unsupervised features extractor through an inverse mapping from data space to the latent space. The reconstruction error is the distance between a data point x and its reconstruction $G(z_\gamma)$ where z_γ is the inverse mapping of x into the latent space.

3 Related Works

Still nowadays the most widely used approaches for anomaly detection in spacecraft telemetries are based on simple Out of Limit (OOL) checks [11], meaning that, when the signal exceeded some predefined upper and lower bounds, an alarm is triggered. More advanced solutions introduced clustering techniques on multidimensional vectors obtained by manually extracted features, like the *Inductive System Health Monitoring* developed at NASA [9] and the *Automated Telemetry Health Monitoring System* (ATHMoS) developed at the German Space Operation Center (GSOC) [16]. Recent works propose the introduction of deep learning. For example [15] proposes the introduction of deep autoencoders to extend the manually extracted features, [12] proposes a deep autoencoder to model the normal behavior of the telemetries and a thresholding technique on the reconstruction errors to detect anomalies and it suggests the introduction of a recurrent architecture to take into account the temporal evolution. In this direction [8] proposes a Long Short-Term Memory (LSTM) recurrent network to predict the future signal under normal conditions, then, at test time, the predicted values are compared with the observed values and anomalies are computed using thresholding techniques. In this paper we propose the application of the GANs framework for the spacecraft telemetries anomaly detection. The GANs have already been used for finding anomalies in complex data. In particular [18] introduced the AnoGAN architecture for the marker discovery in tomography images of the retina. A more efficient version of AnoGAN called Efficient GAN-Based Anomaly Detection (EGBAD) was introduced in [20]. It uses the Bidirectional GANs architecture [4] to learn at training time an inverse mapping from data space to latent space. In this paper we present our implementation of the EGBAD framework along with an analysis of the anomaly scores that employs a histogram for the detection and the temporal evolution of the anomaly score for the localization of the anomaly.

4 Data

We have at our disposal simulated data that emulate the operation of a LEO (Low Earth Orbit) satellite. In particular we studied the sensors monitoring a *Reaction Wheel* (RW), which is a type of flywheel used primarily by spacecrafts for three-axis attitude control. The RW has a high pointing accuracy and it is particularly useful when the spacecraft must be rotated by very small amounts, for example for keeping a telescope pointed at a star. The Reaction Wheel was equipped with four sensors monitoring: the current absorbed,the temperature, the velocity and the commanded torque (see Figure 1a). TAS-Italia industry [1] provided us the simulation of four months of observation data, one of those with an anomalous behavior.



(a) A sample of telemetry data. Each line corresponds to a sensor: motor current (blue line), temperature (orange line) and angular speed (green line)

(b) Graphical representation of the BiGAN architecture. g and d are respectively the generator and the discriminator, while E is the encoder (the novel component that characterizes the BiGAN model). x is a data sample from the data distribution and z' is a sample from the known latent distribution.

Fig. 1

5 Method

In this paper we propose an implementation of the EGBAD framework (represented in Figure 1b). In particular we split the stream of telemetries into fixed length sequences and implemented both the Generator, the Encoder and the Discriminator as multilayer perceptrons. The networks are trained adversarially with the original BiGAN loss. The generator G learns the mapping from samples from the latent distribution $p_z(z)$ to samples from the data distribution $p_x(x)$ while the encoder E learns the inverse mapping. The discriminator D discriminates jointly in data and latent space. Even though it is not explicit, the Encoder and the Generator are proven to be one the inverse of the other at the optimum [4]. The three networks were implemented as reported in table 1. Both in the univariate and multivariate case we

[1] https://www.thalesgroup.com/it/global/activities/space

split the input into fixed length windows. The input shape of the Encoder changes in the univariate and multivariate cases. In the former it equals the sequence length while in the latter it is the sequence length multiplied by four, that is the number of channels, since the input is flattened.

Network	Layer	Units	Non Linearity	Dropout
$E(x)$	Dense	64	Leaky ReLU	0.0
	Dense	64	Linear	0.0
$G(x)$	Dense	64	ReLU	0.0
	Dense	128	ReLU	0.0
	Dense	121	Linear	0.0
$D(x)$	Dense	128	Leaky ReLU	0.2
$D(z)$	Dense	128	Leaky ReLU	0.2
$D(x, z)$	Dense	128	Leaky ReLU	0.2
	Dense	1	Linear	0.0

Table 1: BiGAN architecture.

We trained the model with Adam optimizer with learning rate of $1e - 5$ and betas equal to 0.5. We used a batch size of 512 samples, while the latent dimension was equal to 32. All the weights and biases of the *encoder* and *discriminator* layers are initialized with the $Xavier$ initializer, while the weights and biases of the *generator* are initialized with the He initializer. We trained the network for 5 epochs for each experiments.

Once the model is trained, the anomaly score $A(x)$ is computed as a convex combination of the reconstruction loss L_G and the discriminator-based loss L_D:

$$A(x) = \alpha L_G(x) + (1 - \alpha) L_D(x) \tag{1}$$

where L_G and L_D are defined as follows:

$$L_G(x) = |x - G(E(x))|_1 \tag{2}$$

$$L_D(x) = |f_D(x, E(x)) - f_D(G(E(x)), E(x))|_1 \tag{3}$$

In particular the L_D term is called feature matching loss [21] and f_D is the output of the discriminator layer that precedes the final classification layer. The anomaly score $A(x)$ is not bounded, hence, in order to be interpreted, it must be compared with some reference values, that in our case are the anomalies scores computed on the normal data.

6 Results

The proposed method was introduced to solve a complex anomaly detection task. Figure 2 shows some samples taken from the normal data distribution and from

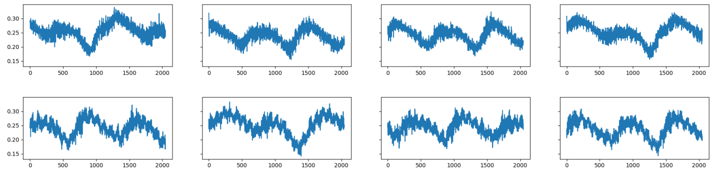


Fig. 2: Sequences sampled from the current sensor, in particular the sequences in the first row are sampled from the normal distribution, while the ones in the second row comes from the anomalous distribution.

the anomalous data distribution. Also from a human analysis the anomaly is not easily identifiable. In such a complex scenario, approaches based on clustering and k nearest neighbour, that are the most frequently mentioned in literature [2], failed to identify the known anomaly. Figure 3 shows the principal components of some points sampled from the normal and the anomalous distributions and it shows that normal and anomalous points are not separable in the feature space. Despite points in figure 3 are just represented with the two principal components we also tried to cluster normal versus anomalous points using more than two principal components, but the results remained almost unchanged. We also tried density estimation with Dynamic Time Warp (DTW): we fixed a threshold for the DTW distance and for each sample we counted the number of normal neighbours (with the DTW below the threshold). We tried different thresholds but the expected number of neighbours was approximately the same both for normal and anomalous samples. Figure 4 shows the 1-NN (taken from the normal samples) of a normal and an anomalous sample and the respective warping curve. We can see that they are very similar and the anomalous vs normal nearest neighbour warping curve is actually closer to a straight line with respect to the normal vs normal nearest neighbour warping curve. The failures of the classical approaches led us to the proposed method based on Generative Adversarial Networks, which results will be detailed in the rest of this section. In particular we

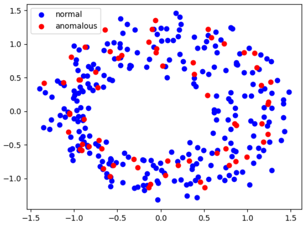


Fig. 3: First two principal components of the data samples (from the current sensor). The blue and red points represent respectively the normal and anomalous data points.

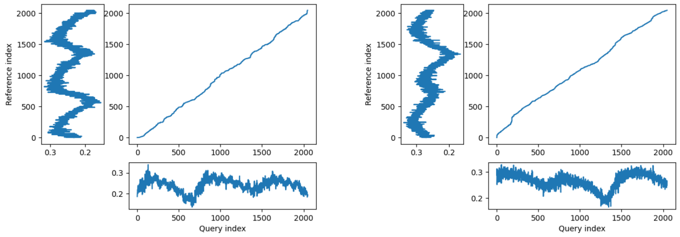
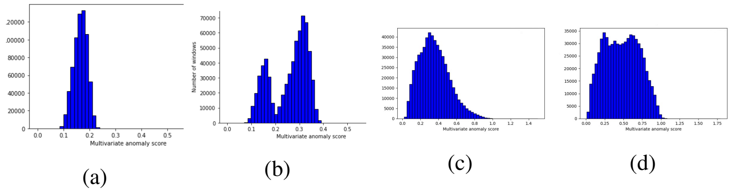


Fig. 4: Warping curve between the query sample (at the bottom of the curve) and its closes neighbours in the normal samples (at the left of the curve). In the left figure the query is a normal sample, while in the right figure the query is an anomalous sample. In both cases the query refers to the current sensor.

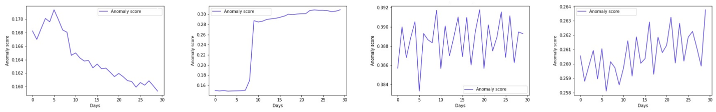
carried both a univariate and multivariate analysis. In the former case the network is trained independently on each sensor, therefore the correlation between sensors is not considered. In the latter the network is trained once for all sensors, hence the input of the network is a multivariate time series in which each channel corresponds to a particular sensor. In both configurations we analyzed the anomaly scores varying the temporal granularity. We compared two observation periods that were not used for training. The first one, referred in the following as *Normal Period*, contains only normal data while the second, referred as *Anomalous Period* contains a known anomaly. For the sake of clarity let us specify that not every sample in the anomalous period is affected by an anomaly but rather there is just a single anomaly and its time localization is known. First we plotted an histogram of the anomaly score for each period: the range of the anomaly score is split into bins on the horizontal axis and the bars on the vertical axis are proportional to the frequency of each bin. The histograms of the two periods for the univariate and multivariate case are



(a) (b) (c) (d)

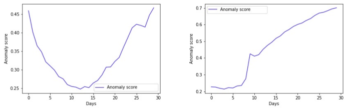
Fig. 5: Histograms counting the frequencies of the anomaly scores in a given observation period. In particular the first column (figures 5a and 5c) refers to the period without anomalies while the second column (figures 5b and 5d) to the one with the known anomaly. The First row (figures 5a and 5b) shows the histograms for the univariate case and in particular they refer to the current sensor, while the second row (figures 5c and 5d) illustrates the histogram for the multivariate case.

depicted in Figure 5. It is important to notice that the anomaly is detected by both the univariate and multivariate approaches even though the plots are significantly different. In particular in the univariate case (top row in the Figure 5) the curves outlined by the histograms have the shape of a unimodal distribution for the normal period and a bimodal distribution for the anomalous period. The peaks correspond to the anomaly scores with higher frequency. We can observe that the smaller peak in the anomalous period and the one in the normal period correspond to the same amount of anomaly score, hence they can be interpreted as the normal samples, while the bigger peak in the anomalous period matches a larger anomaly score, therefore it is probably caused by the anomalous sequences. The picture is different for the multivariate case in which we have unimodal distributions in both the normal and anomalous observation period. However the latter has higher mean and standard deviation, which are both indications of abnormal behavior. Although these graphs allow to individualize an anomaly they do not allow to individualize the instant, within the observation period, in which the anomaly has occurred. At this purpose we also plotted the evolution of the anomaly score with respect to time for both the normal and anomalous periods. It worth noting that the plots corresponding to different sensors do not always agree. For example Figure 6 represents the evolution of the anomaly score for the current and speed sensors. In the former plot we can see a pulse that can be associated with the known anomaly while the latter is very noisy and it does not allow to identify an anomalous event. In the multivariate case (Figure 7) the plot is more noisy, since some sensors were affected by the anomaly



(a) (b) (c) (d)

Fig. 6: The figures show the temporal evolution of the anomaly score in different observation periods and different sensors. In particular figures 6a and 6b refers to the current sensor while figures 6c and 6d to the speed sensor. Figures 6a and 6c refers to the normal period while figures 6b and 6d to the anomalous one.



(a) (b)

Fig. 7: The figures show the temporal evolution of the anomaly score in the multivariate case for the normal (Figure 7a) and anomalous (Figure 7b) periods.

while others do not. Also in this case however there is a sudden increment in the anomaly score that corresponds to the known anomaly.

7 Conclusion

In this paper we applied the Generative Adversarial Networks, and in particular the BiGAN architecture, to the problem of anomaly detection in spacecraft telemetry data and compared the results with those of some classical algorithms. In our experiment the proposed method was the only one capable of identifying the known anomaly. Moreover the complexity of the model with respect to the dataset is drastically reduce with respect to methods based on KNN, since once the model is trained, the execution is very fast and do not depends on the size of the dataset. From the univariate and multivariate analisys emerged that there is no strong correlation between sensors and in particular some sensors seem not to be affected by the anomaly, while others showed a sudden increment of the anomaly score when the anomaly is happening. This fact is positive from an implementation point of view because the model scales well with the input. In particular we considered only four sensors connected to a reaction wheel but the system could be easily expanded for the monitoring of all the sensors in the spacecraft. This is not the case in the multivariate case, because the number of parameters is quadratic with respect to the input size. It could be the case however that the identification of more complex anomalies could depend on the complex relations between different sensor readings. At this purpose it would be interesting for a future work to test the model with more anomalies and different configurations of the input sources.

References

[1] Varun Chandola, Arindam Banerjee, and Vipin Kumar. "Anomaly Detection: A Survey". In: *ACM Computing Surveys* 41.3 (July 2009), pp. 1–72. ISSN: 0360-0300. DOI: 10.1145/1541880.1541882. URL: https://doi.org/10.1145/1541880.1541882.

[2] Varun Chandola, Arindam Banerjee, and Vipin Kumar. "Anomaly detection: A survey". In: *ACM computing surveys (CSUR)* 41.3 (2009), pp. 1–58.

[3] Meimei Ding and Hui Tian. "PCA-based network traffic anomaly detection". In: *Tsinghua Science and Technology* 21.5 (2016), pp. 500–509.

[4] Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. "Adversarial feature learning". In: *arXiv preprint arXiv:1605.09782* (2016).

[5] Ian Goodfellow et al. "Generative Adversarial Networks". In: *Commun. ACM* 63.11 (Oct. 2020), pp. 139–144. ISSN: 0001-0782. DOI: 10.1145/3422622. URL: https://doi.org/10.1145/3422622.

[6] Xiaoyi Gu, Leman Akoglu, and Alessandro Rinaldo. “Statistical analysis of nearest neighbor methods for anomaly detection”. In: *Advances in Neural Information Processing Systems* 32 (2019).

[7] Ling Huang et al. “In-network PCA and anomaly detection”. In: *Advances in neural information processing systems* 19 (2006).

[8] Kyle Hundman et al. “Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding”. In: *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 2018, pp. 387–395.

[9] David L Iverson. “Inductive System Health Monitoring.” In: *IC-AI*. 2004, pp. 605–611.

[10] Istvan Kiss et al. “Data clustering-based anomaly detection in industrial control systems”. In: *2014 IEEE 10th International Conference on Intelligent Computer Communication and Processing (ICCP)*. IEEE. 2014, pp. 275–281.

[11] Jose Martinez. “New telemetry monitoring paradigm with novelty detection”. In: *SpaceOps 2012*. 2012, p. 1275123.

[12] Jose Martinez and Alessandro Donati. “Novelty Detection with Deep Learning”. In: *2018 SpaceOps Conference*. 2018, p. 2560.

[13] Gerhard Münz, Sa Li, and Georg Carle. “Traffic anomaly detection using k-means clustering”. In: *GI/ITG Workshop MMBnet*. Vol. 7. 2007, p. 9.

[14] Christian Napoli et al. “Exploiting Wavelet Recurrent Neural Networks for satellite telemetry data modeling, prediction and control”. In: *Expert Systems with Applications* (2022), p. 117831.

[15] Corey OMeara, Leonard Schlag, and Martin Wickler. “Applications of deep learning neural networks to satellite telemetry monitoring”. In: *2018 SpaceOps Conference*. 2018, p. 2558.

[16] Corey OMeara et al. “ATHMoS: Automated telemetry health monitoring system at GSOC using outlier detection and supervised machine learning”. In: *14th International Conference on Space Operations*. 2016, p. 2347.

[17] Alec Radford, Luke Metz, and Soumith Chintala. “Unsupervised representation learning with deep convolutional generative adversarial networks”. In: *arXiv preprint arXiv:1511.06434* (2015).

[18] Thomas Schlegl et al. *Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery*. 2017. arXiv: `1703.05921 [cs.CV]`. URL: `https://arxiv.org/abs/1703.05921`.

[19] Iwan Syarif, Adam Prugel-Bennett, and Gary Wills. “Unsupervised clustering approach for network anomaly detection”. In: *International conference on networked digital technologies*. Springer. 2012, pp. 135–145.

[20] Houssam Zenati et al. “Efficient gan-based anomaly detection”. In: *arXiv preprint arXiv:1802.06222* (2018).

[21] Yizhe Zhang et al. “Adversarial feature matching for text generation”. In: *International Conference on Machine Learning*. PMLR. 2017, pp. 4006–4015.