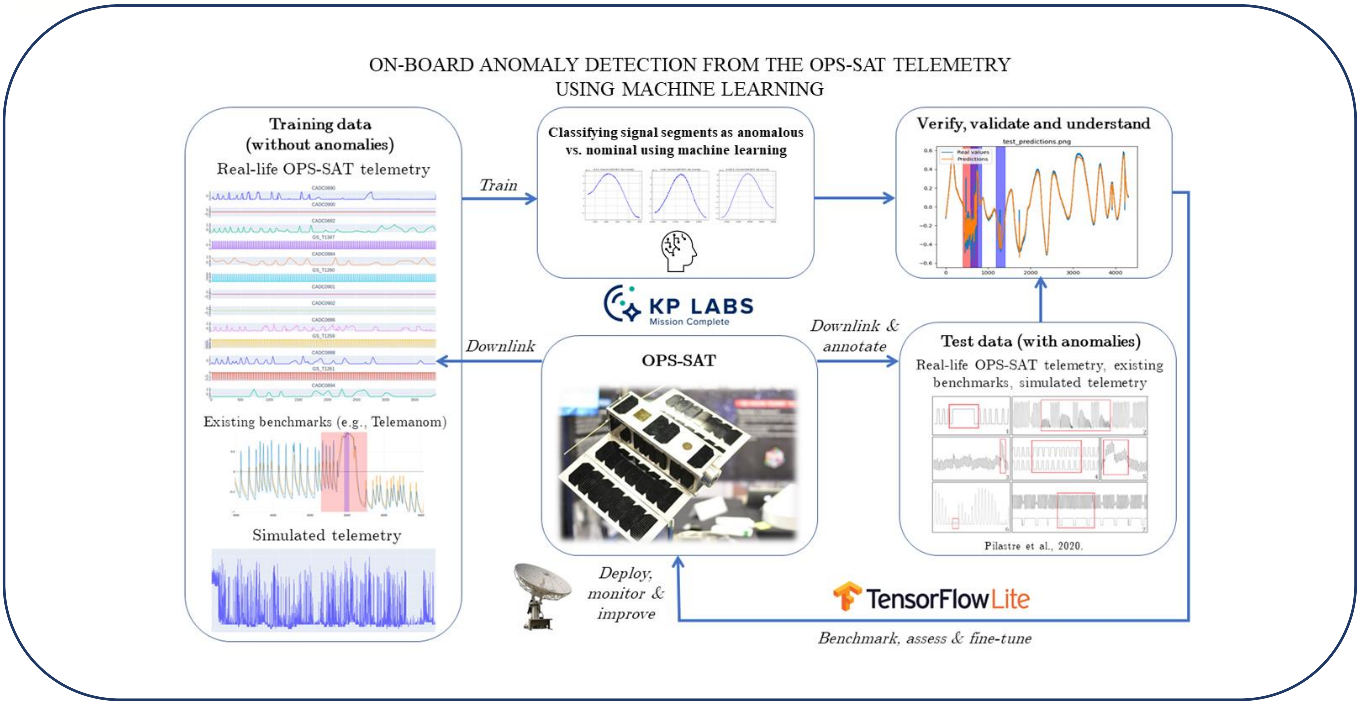

On-board anomaly detection from the OPS-SAT telemetry using deep learning

Executive Summary Report

Study

OSIP OPS-SAT EXPERIMENTS CAMPAIGN - STUDIES
Affiliation(s): KP Labs Sp. z o. o.

Activity summary:

Detecting anomalies from the satellite telemetry is critical for its safe operation. Although many approaches to autonomous on-board anomaly detection have already been proposed, most of them have so far only been tested on non-satellite or simulated data. We tackled this research gap and proposed an end-to-end machine learning-powered approach for detecting abnormal events in real-life OPS-SAT telemetry data, and deployed it on board a satellite. The experimental study revealed that our technique achieves the classification accuracy of 0.957 over the unseen & validated test set, with precision and recall of 0.929 and 0.897, respectively, while offering very fast inference.

Publishing Date: 12.12.2022
Contract Number: 40001373339/22/NL/GLC/ov
Implemented as ESA Express Procurement Plus (EXPRO+)

Contents

1 Introduction

1.1 Purpose and scope

This document summarizes the outcomes and activities executed under the Contract On-board anomaly detection from OPS-SAT telemetry using deep learning, ESA Contract No. 40001373339/22/NL/GLC/ov. It is a self-standing document, suitable for non-technical audience, containing essential information, relevant for each Work Package of the Project.

1.2 Applicable documents

The following documents are applicable to this report as they contain relevant information:

1. ESA Contract No. 40001373339/22/NL/GLC/ov
2. Idea_I-2021-03771_2nd_Round_On-board_
3. OSIP Cover letter for proposal_2-1816
4. OPS-SAT Progress Report
5. OPS-SAT Final Report

2 Project Summary

2.1 WP100 – Datasets

The dataset of telemetries consisting of real-life OPS-SAT telemetry signals has been prepared for this work package. The set of signals has been selected following the advice of OPS-SAT operators, who pointed out some telemetries as interesting. Thus, for the described dataset, several telemetry signals from the magnetometer (channel names: CADC0872, CADC0873, CADC0874) and the PD channels (channel names: CADC0884, CADC0886, CADC0888, CADC0890, CADC0892, CADC0894) have been considered. The resulting dataset contained 524 anomaly candidates, and after the revision, it has 445 identified anomalies of different types, including single peak anomalies, anomalies occurring on a single channel, and anomalies occurring on various channels. The data source of the gathered signals was the MUST, which is a system for the OPS-SAT telemetric data access. We exploited three different augmentation approaches to synthesize additional nominal training data that can be used to train better generalizing machine learning models. Those techniques included mirroring the signal over both axes, and shifting the signal by a certain number of steps. In Figure 1, we show several examples of the signal segments from our dataset, with the indication if they are labelled as anomalous or not.

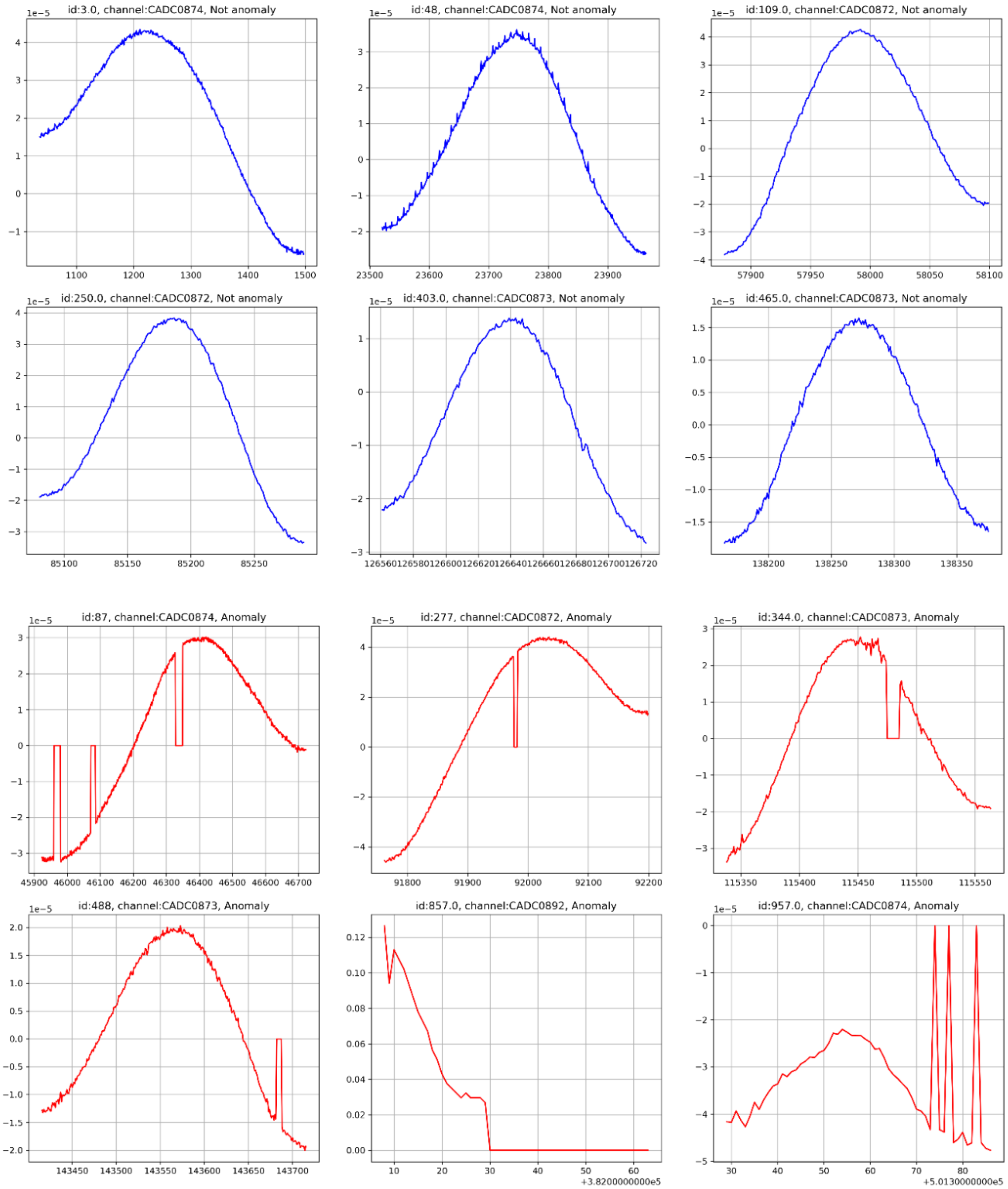

Figure 1. Examples of nominal segments (in blue) and segments labeled as anomalous (red) of several analyzed telemetry channels.

2.2 WP200 – Algorithms

To detect anomalous events in OPS-SAT telemetry signals, we utilized a classic machine learning pipeline utilizing a random forest model operating over the features extracted from the signal's segments. We benefit from almost twenty feature extractors which can effectively process segments of various lengths, including those extracting basic statistics of the segments (their duration, length, mean, variance, standard deviation), the number of peaks (for the original and smoothed segments, as well as for the segment's first and second derivatives), the variance of the segment's first and second derivative, squared number of missing readouts, the weighted segment length (relative to its sampling), and the variance relative to segment's length and duration. The final classification model has been prepared using the refined version of the training set, built of the verified labelling for anomalous and non-anomalous samples, and included augmented data, achieving the classification accuracy of 0.957 over the test set, with the precision and recall of 0.929 and 0.897, respectively. The qualitative analysis revealed that the incorrectly classified segments, e.g., false positives, would be challenging to the human operators too, as they manifest only subtle characteristics which may indicate that they are abnormal. Additionally, the most challenging segments include those that are positioned at the "boundary" of the anomalous and nominal parts of the telemetry channel. The resulting model has also been investigated if any of the features from the dataset contribute with a greater importance to the final classification.

2.3 WP300 – Software

The main purpose of our satellite software is to acquire data, perform experiments and store their results. To do this, we used NanoSat MO Framework to make our application able to communicate with the rest of the satellite systems, and perform data acquisition process. Next, we perform data preprocessing, which consists of data resampling and splitting it into batches used later in the detection process. The software part also reimplements the algorithms provided in WP200 to be able to run them in more demanding and constrained execution environment. Also, the flight software must be stable and tested, so the available libraries are often thoroughly validated but they are in older or even legacy versions, and some of widely-used libraries are not available at all. All those things must be considered while implementing software for the satellite. Our anomaly detection algorithm is not memory demanding and, for the CPU, our benchmarks shows that we spend more time on data acquisition process than on the actual inference. That means that our solution can perform detections without dropping any samples and (depending on preprocessing and batch size) perform near real time operations which is pivotal when it comes to anomaly detection. Finally, the experimental results can be downloaded according to their priority: abnormal segments are considered high priority and those which are nominal, as a low priority. There is an easy access to those results within the OPS-SAT internal systems for further analysis.

2.4 WP400 – Research management and dissemination

The results of the projects were presented at the 8th International Workshop on On-board Payload Data Compression (OBPDC 2022) conference in Athens, Greece (28-30 September 2022), see https://opssat1.esoc.esa.int/attachments/download/870/Publication_OBPDC%202022t.pdf.
Currently, we are working on a journal paper summarizing our approach toward detecting anomalous events from telemetry channels based on classic machine learning approaches and feature extractors operating on the signal segments of any length.