Space Imagesets Don’t Exist!

The defining problem of AI/CV in space

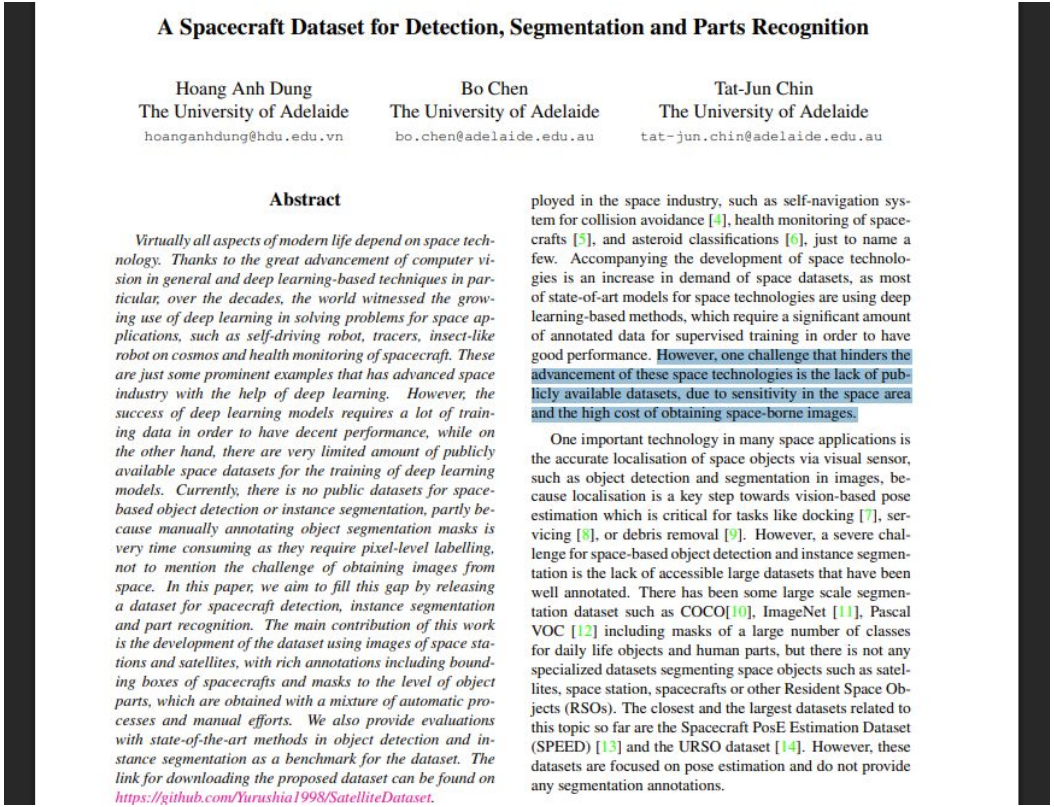
Problem

Spacecraft Images

On this page, you will find a collection of MAVEN spacecraft images. Click on each image for a full resolution version.

Artist renditions

Problem



A Spacecraft Dataset for Detection, Segmentation and Parts Recognition

Hoang Anh Dung
The University of Adelaide
hoanganhdung@hdu.edu.vn

Bo Chen
The University of Adelaide
bo.chen@adelaide.edu.au

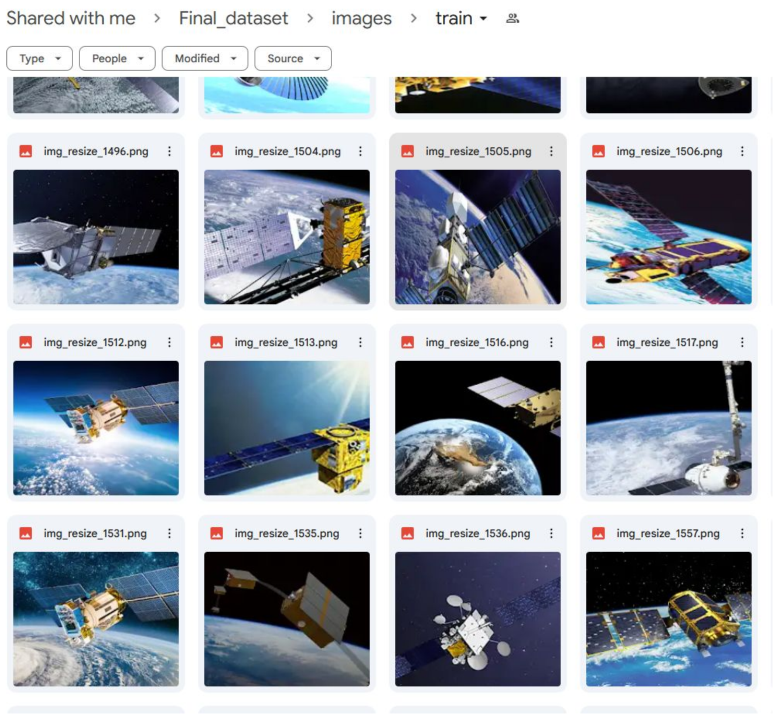
Tat-Jun Chin
The University of Adelaide
tat-jun.chin@adelaide.edu.au

Abstract

Virtually all aspects of modern life depend on space technology. Thanks to the great advancement of computer vision in general and deep learning-based techniques in particular, over the decades, the world witnessed the growing use of deep learning in solving problems for space applications, such as self-driving robot, tracers, insect-like robot on cosmos and health monitoring of spacecraft. These are just some prominent examples that has advanced space industry with the help of deep learning. However, the success of deep learning models requires a lot of training data in order to have decent performance, while on the other hand, there are very limited amount of publicly available space datasets for the training of deep learning models. Currently, there is no public datasets for space-based object detection or instance segmentation, partly because manually annotating object segmentation masks is very time consuming as they require pixel-level labelling, not to mention the challenge of obtaining images from space. In this paper, we aim to fill this gap by releasing a dataset for spacecraft detection, instance segmentation and part recognition. The main contribution of this work is the development of the dataset using images of space stations and satellites, with rich annotations including bounding boxes of spacecrafts and masks to the level of object parts, which are obtained with a mixture of automatic processes and manual efforts. We also provide evaluations with state-of-the-art methods in object detection and instance segmentation as a benchmark for the dataset. The link for downloading the proposed dataset can be found on https://github.com/Yurushia1998/SatelliteDataset.

ployed in the space industry, such as self-navigation system for collision avoidance [4], health monitoring of spacecrafts [5], and asteroid classifications [6], just to name a few. Accompanying the development of space technologies is an increase in demand of space datasets, as most of state-of-art models for space technologies are using deep learning-based methods, which require a significant amount of annotated data for supervised training in order to have good performance. However, one challenge that hinders the advancement of these space technologies is the lack of publicly available datasets, due to sensitivity in the space area and the high cost of obtaining space-borne images.

One important technology in many space applications is the accurate localisation of space objects via visual sensor, such as object detection and segmentation in images, because localisation is a key step towards vision-based pose estimation which is critical for tasks like docking [7], servicing [8], or debris removal [9]. However, a severe challenge for space-based object detection and instance segmentation is the lack of accessible large datasets that have been well annotated. There has been some large scale segmentation dataset such as COCO[10], ImageNet [11], Pascal VOC [12] including masks of a large number of classes for daily life objects and human parts, but there is not any specialized datasets segmenting space objects such as satellites, space station, spacecrafts or other Resident Space Objects (RSOs). The closest and the largest datasets related to this topic so far are the Spacecraft PosE Estimation Dataset (SPEED) [13] and the URSO dataset [14]. However, these datasets are focused on pose estimation and do not provide any segmentation annotations.



Problem

- AI has an endless appetite for images
- ~All imagesets of space objects are not real!
- As AI+space matters more, so does this gap
- Labeling, curating existing sources is clunky

Why this gap exists

- Hard to get cameras into space as-is
- Expensive to downlink images en masse
- “Sensitivity” aka vague legal reasons
- Cameras are often looking at ~nothing in space
- Simple visual environment; easy to render instead
- Takes a lot of work to label those images well

What non-synthetic data does exist

- Not much is public
- Private rendezvous imagery
- Miscellaneous, unlabelled videos
- SPEED, SPEED+, Stanford etc

Really very little!

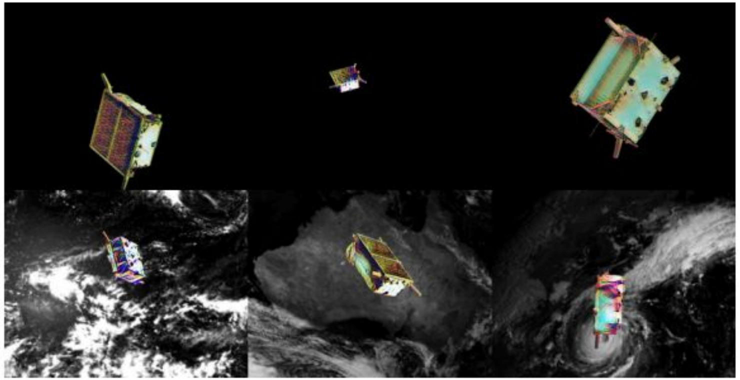


Figure 6. Examples of 6 images from PRISMA12K-TR

inference, this NST pipeline allows for randomly sampling a vector of style embedding $z \in \mathcal{R}^{100}$. Specifically, the style embedding is sampled as

$$z = \alpha \mathcal{N}(\mu, \Sigma) + (1 - \alpha) P(c) \tag{10}$$

where $P(c)$ is the style embedding of the content image, (μ, Σ) are the mean vector and covariance matrix of the style image embeddings pre-trained on ImageNet, and α is the strength of the random normal sampling. In this work, $\alpha = 0.25$ is used to create PRISMA12K-TR. In order to avoid the NST's blurring effect on the spacecraft's edges, the style-randomized spacecraft is first isolated from the background using a bitmask then combined with the original background. Figure 6 shows a montage of six such images.

The third dataset is PRISMA25, which consists of 25 spaceborne images captured during the rendezvous phase of the PRISMA mission.[33] The PRISMA25 is used to evaluate the performance of the CNN on a previously unseen spaceborne dataset when trained solely on a mixture of PRISMA12K and PRISMA12K-TR.

More specific challenges

- We can always generate synthetic data, but does it validate?
- Real imagery contains… lighting aberrance, cluttered field, edge cases!
- Little standardization in camera models