Onboard-Targeted Segmentation of Broken Pixels, Lines, and Dust on Optics Faults in Space Camera Sensors

Anonymous CVPR submission

Paper ID *****

Abstract

This study proposes a deep learning framework for the semantic segmentation of faults in space camera sensors, specifically addressing dust grains deposition on the optics and defect pixels and lines caused by radiation impinging the detector. The artifacts caused by these faults in the final image pose significant challenges in the field of thin structures segmentation. To address this, the DeepLabV3 architecture is employed, with various configurations evaluated to improve performance. A MobileNetV3 backbone is utilized to target compatibility with onboard deployment on resource-constrained hardware. Furthermore, an integration strategy to operate the model within the onboard navigation pipeline is proposed, highlighting its utility in improving autonomy and availability of the mission. These improvements are evaluated by means of integration-aware metrics in order to quantify the functional benefits of the proposed model.

1. Introduction

The increasing reliance on camera sensors of the navigation systems of modern space missions poses significant challenges for Fault Detection, Isolation and Recovery (FDIR). In addition, future planetary exploration missions require an increasing level of autonomy of the spacecraft functionalities, which motivates the need of onboard, autonomous FDIR. AI-based FDIR solutions can tackle the gaps of traditional FDIR based on ECSS-E-ST-70-41C (Packet Utilization Standard, [14]), enabling autonomous and potentially real-time image processing. The aim for reliable, embedded-compatible AI solutions motivates the present research.

The technological demonstrator mission [29] aims to perform autonomous relocations on the surface of the 67P/Churyumov-Gerasimenko comet, utilizing AI-augmented Guidance [31], Navigation [24], and FDIR subsystems [16, 18]. These subsystems integrate LiDAR, camera sensors, gyroscopes, and accelerometers to accomplish mission goals. The biggest challenge of the FDIR task in camera data is the need to identify and categorize failure cases for this class of sensors, detailing their applicability to different mission scenarios. While many interplanetary exploration missions consistently employ these sensors [2, 3, 32, 37, 46–48], the field of space camera faults is rather unexplored in the state of the art. The same consideration applies to research in the commercial cameras domain. Typically, deteriorating effects in space cameras are treated with traditional software corrections or hardware solutions. For instance, the Cassini mission Imaging Science System (ISS) addresses broken pixels due to radiation effects on the camera sensors with *a posteriori* calibration [32, 46]. Furthermore, dust deposition on the optics of the Curiosity rover is mitigated via a proper re-closable mechanical dust cover, which is only opened when the camera is operational [48].

A prior failure mode analysis [17] resulted in the release of an open-access dataset specifically curated for training and testing AI algorithms for FDIR. Artifacts derived from camera faults can obstruct the sensor's Field of View (FoV) and introduce anomalies during data readout. When processed by onboard software, such corrupted data may degrade performance or induce errors in nominal operations. Building upon the foundational characterization in [17], the current study addresses fault detection through AI-based semantic segmentation. Our solution shall aim to achieve fail-operational conditions where the AI algorithm segments faulty pixels and allows to exclude them from successive processing of images, thus preventing failure propagation while maintaining system operability. The specific safety-critical FDIR framework requires increased reliability and thorough analysis of the functionality as integrated alongside other components of the spacecraft. Moreover, evaluating the effect that incorrect AI-triggered fault detection may have at subsystem or system level shall be part of the development phase, with the goal of deriving performance indices and requirements definition.

This work proposes an onboard-compatible deep learn-

ing solution for the semantic segmentation of Dust on optics and Broken Pixels and Lines. The performance of this architecture is evaluated using novel, system-aware metrics designed to reflect the requirements of downstream processing operations. The selected AI model is tuned for the fault segmentation task, initially targeting all faulty classes within the chosen dataset [17], subsequently focusing exclusively on the Dust on Optics and Broken Pixels and Lines. During the design process, challenges associated with the segmentation of thin-structured artifacts resulting from the targeted faults were encountered; these were efficiently addressed through an iterative analysis leveraging variations in the model's architectural blocks. The selected model prioritizes onboard deployment compatibility, utilizing embedded-friendly components to minimize the computational effort and memory footprint on the spacecraft hardware. Finally, we propose a strategy for interfacing the semantic segmentation model with the navigation subsystem, specifically focusing on how it can leverage the fault information derived from the AI prediction. Building upon this operational context, we derive custom metrics to accurately measure the model's performance as integrated into the described framework.

This paper is structured as follows: Section 2 discusses the related work in the field of embedded-compatible semantic segmentation, its application in space and in thin structures analysis. Section 3 describes the development process we followed to get to our AI-based segmentation solution for camera faults. Section 4 discusses the best model solution derived from the analysis, the evaluation metrics and some considerations about the model integration onboard. Finally, Section 5 presents the conclusions and future work.

2. Related Work

Semantic segmentation refers to the task of performing dense pixel-level predictions. State-of-the-art architectures for semantic segmentation of images employ fully convolutional layers to output segmentation masks [20, 27, 33, 39]. Recently, the use of Transformers [43] shifted the focus from convolutions to this new architecture, which is proven to work well in segmentation [11, 26], as well in many other tasks, e.g., Natural Language Processing.

The problem of applying segmentation architecture on non-optimized and computationally-limited hardware, i.e., CPUs and mobile phones, has been thoroughly investigated [4]. Depthwise separable convolutions were firstly employed in the Inception models [39], Xception network [8] and successively by the MobileNet series [21, 22, 34], which targeted efficient inference in mobile phones and embedded devices. In [34], the authors firstly introduced the *Inverted Residual Block* and proved its memory efficiency due to the residual connection between the bottlenecks. Another approach to model embedded-compatible architectures was proposed in the EfficientNet series [40, 41], which leverage Neural Architecture Search (NAS) to propose a family of models with fewer parameters and faster inference than state-of-the-art models. Specifically, [41] makes use of the Fused-MBConv block, a modified version of the Inverted Residual Block. *Dilated Convolutions* (or *Atrous*) were introduced in [49], to enlarge the receptive field of the convolution without increasing the number of operations or memory footprint, with respect to conventional convolutions. Dilated Convolutions are employed in compact models such as MobileNetV3 [21] and ESPNet [30]. Focusing more on the semantic segmentation task, Spatial Pyramid Pooling (SPP) is proposed for multi-scale context aggregation [19]. This method allows learning of features at different scales by concatenating feature maps obtained through different downsampling rates. In addition, SPP also allows to obtain a fixed-size feature representation, removing any requirement on the size of the network input image. Pioneering work in the SPP field is the DeepLab series [5–7], which introduced the Atrous Spatial Pyramid Pooling (ASPP) module, leveraging Atrous convolutions for larger-scale feature learning at competitive computational cost. ASPP is applied on the low-dimensional feature output of the backbone, while bilinear upsampling serves to retrieve the original input image size. In DeepLabV3+ [7], this non-learnable upsampling was deemed responsible to miss important segmentation features and lead to inaccurate output, especially on small objects. To solve this issue, the author concatenated low-level features from the convolutional backbone with the output of the ASPP and processed with depthwise separable convolutions, substituting the bilinear upsampling with a learnable module. Alternatively, [45] proposed Dense Upsampling Convolution (DUC), providing a learnable upsampling module based on dense convolutions to retrieve segmentation masks from downsampled backbone outputs. Further Work on segmentation of small objects concentrates on data augmentation [23], leveraging on the one hand a scaling and stretching approach, on the other hand proposing to artificially duplicate small objects in images to provide more samples to learn.

Embedded AI applications are typically benchmarked on standard datasets [9, 10, 12, 25] using Commercial-off-the-shelf (COTS) hardware. In the aerospace domain, lightweight AI applications specifically targeting onboard constraints are discussed in [1], while space-qualified hardware for AI processing is explored in [15]. Further applications targeting such hardware are detailed in [13], with recent advancements including the deployment of GPUs in space environments [35]. Despite these advancements, literature regarding anomaly detection specifically for spaceborne camera sensors remains limited. Notable examples include AI-based anomaly detection for remote sensing [44]

and the characterization of a dataset of faulty camera images for space applications [17]. Both these works explore failure modes in Charge-coupled Device (CCD) sensors, which are addressed in this work.

3. Methodology

3.1. Model Search

The present analysis targets onboard deployment leveraging space hardware-compatible designs. However, actual hardware validation and deployment are deferred to future work.

To minimize the computational complexity (FLOPs) and the memory footprint during inference, Atrous convolutions are prioritized over conventional convolutions. The requirement for a reduced memory footprint leads to the exclusion of any model employing skip connections on high-dimensional features (e.g., *U-Net*), as these must be retained in memory during the forward pass. Nevertheless, employing the ASPP module [6] allows to exploit the memory benefit of Atrous convolutions while providing enough multi-scale context to enable skip connections avoidance. The *DeepLabV3* architecture is selected based on these requirements. The original implementation of [6] employs *ResNet-50* and *ResNet-101* [20] as backbone models to provide the intermediate feature maps to the DeepLab head. These models are very memory-hungry, due to their depth and the presence of residual connections in each block. In the present work we substitute ResNet with *MobileNetV3 Large* for its high inference speed on edge devices, especially mobile phones. The target hardware of MobileNet is deemed representative for initial qualification of this model to space hardware. Based on the downsampling rate of MobileNet and the use of Atrous convolutions, it is possible to set the $output_stride$ parameter in DeepLabV3 to obtain lower dimensional features handling and consequently reduce the memory footprint of the segmentation head. Note that this choice of $output_stride$ is possible due to the enlarged context aggregation provided by Atrous convolutions of MobileNet. Conventional convolutional nets, such as ResNet employed in the original work, are proven to perform better with $output_stride = 8$ [6], as the downsampling is reduced. Finally, the Atrous rates in the DeepLab head are kept to (6,12,18), as suggested by the optimal results of [6].

3.2. Development Process

Dataset The present analysis employs the specialized dataset from [17] for the fault segmentation task. The dataset comprises a total of five fault classes: *Dust on Optics*, *Broken Pixels and Lines*, *Straylight*, *Optics Degradation* and *Vignetting*. The blurriness and vignetting are labeled with a binary index, because they affect the whole image, so they cannot be addressed through segmentation. The vignetting effect is not critical as its intensity is usually not enough to significantly affect the visibility. Instead, it can be seen as a disturbance effect in estimating and augmenting the information provided by the single pixel. On the contrary, while blurriness can be critical, it occurs on a much longer time scale with respect to the other faults considered. We decide to retain both vignetting and blurriness in the training dataset, assuming that the results of the present analysis will be robust to these two effects, leaving to future research the task of estimating their intensity and embed this information in the overall FDIR pipeline.

Table 1. Dataset properties

Property	Value
Number of Images	2000
Training/Validation Split	80/20
Input Size	$1024 \times 1024 \times 3$
Color Bit Depth	8 bits per channel
Number of Classes	2 (*Nominal*, *Faulty*)

The specialized dataset employed in this work is summarized in Table 1.

Faults Clustering Initial experiments were conducted by training the model on the full dataset, performing binary segmentation of faults where annotation is available (i.e., Dust on Optics, Broken Pixels and Lines, Straylight). Note that *Broken Pixels and Lines* relate to three different failure modes, namely *Broken CCD Pixels*, *Broken Complementary Metal-Oxide Semiconductor (CMOS) Pixels* and *Broken Lines*. The results from the DeepLabV3 model [6] showed good performances on the flares, whereas Dust grains and Broken artifacts were not correctly segmented (Figure 1). The behavior depicted in Figure 1 also highlights that the model is not capable to accurately segment small and thin structures, especially predicting joint artifacts when they are close but separated by a few pixels. Moreover, the segmented area is usually bigger than the underlying grain, pixel or line, suggesting a reduced accuracy at small scales.

Subsequently, we decide to train three different models for Straylight, Broken Pixels and Dust on Optics faults. We adopt this strategy aiming to confirm that one model is not powerful enough to grasp features from all the considered faults and eventually find a tailored design for each fault.

Preliminary Results The separate instances of the standard DeepLabV3 model are trained to perform binary segmentation of broken pixels artifacts and dust grains separately. We adopt a weighted binary cross-entropy loss func-

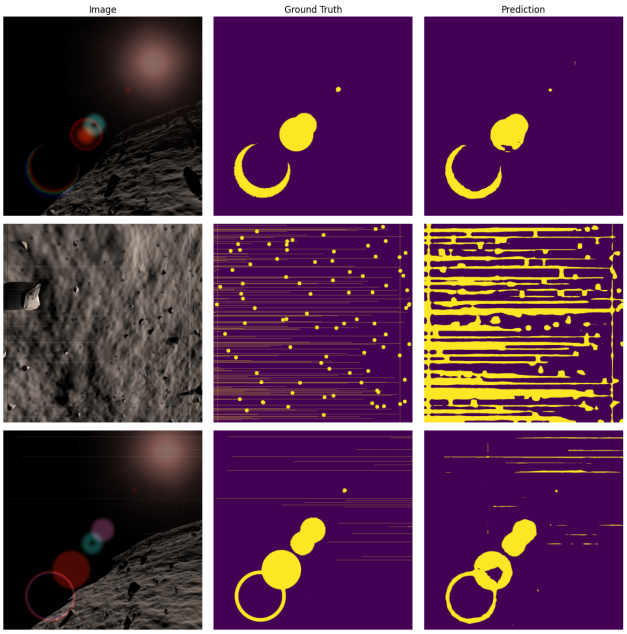

Figure 1. Binary segmentation of all fault cases with DeepLabV3. We train on data presented in Table 1 for 500 epochs and batch size equal to 16, employing Binary Cross-entropy loss and Adam optimizer with 0.0001 initial learning rate.

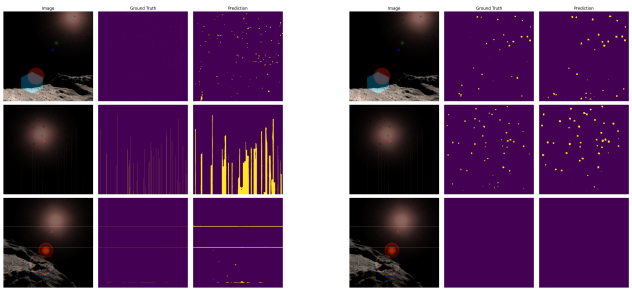

Figure 2. Binary segmentation of Broken Pixels and Lines (left) and Dust on Optics (right) with standard DeepLabV3 model.

tion during training to supply to the class imbalance, where the weight is computed as the faulty/all pixels ratio. We train the models for 500 epochs and batch size equal to 16. The results are shown in Figure 2.

Figure 2 highlights a general improvement of segmentation performances with respect to Figure 1, proving that specialized models perform better over the specific data distributions of faulty pixels. However, poor accuracy in segmenting thin structures is again observed, particularly in the area covered by broken lines and pixels. This issue is also observed in dust grains, albeit less prominently. We adopt two different strategies, supported by state-of-the-art findings (Section 2), to address these issues.

First, as remarked in the original DeepLabV3 work, the

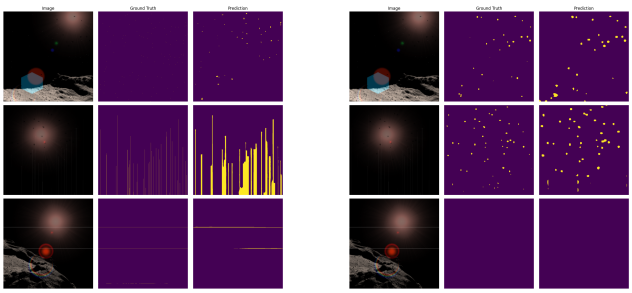

Figure 3. Binary segmentation of Broken Pixels and Lines (left) and Dust on Optics (right) with Atrous rates equal to (6,12,18) in the ASPP module.

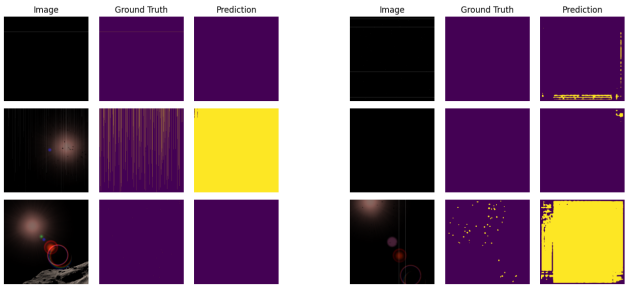

Figure 4. Binary segmentation of Broken Pixels and Lines (left) and Dust on Optics (right) with $ouput_stride = 8$.

ASPP rates provide poor performances when the size of the features to segment is too small. The Atrous convolutions would mostly encounter zeros in the computations and would not be effective in their long-scale-context learning. Hence, we decrease the Atrous rates in the ASPP module to (6,12,18).

Second, the performance degradation can also be caused by an excessive downsampling rate applied in the backbone network, which cannot retain the information about thin artifacts in its deeper layers. In this regard, it is possible to change the $output_stride$ parameter in the DeepLabV3 model to let the ASPP module target an upper layer of the MobileNet, such that the downsampled features are $1/8$ of the input size. Due to the lower downsampling rate, more information on the represented artifacts should be included in the backbone output and consequently seen in the segmentation head.

The results of both the approaches are shown in Figure 3 and Figure 4.

The introduction of reduced Atrous rates in the ASPP module (Figure 3) is shown to be mildly effective in segmentation performances with respect to the standard model version (Figure 2). The two models were run on the same three images and yielded nearly identical segmentation masks; any differences appear to stem from stochastic factors rather than actual performance differences.

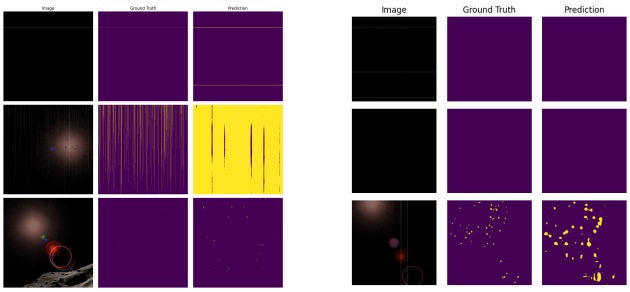

Figure 5. Binary segmentation of Broken Pixels and Lines (left) and Dust on Optics (right) with stretch and rescale pre-processing.

Figure 4 provides evidence that acting on downsampling does not affect the thin structures segmentation problems. Conversely, a reduced $output_stride$ increasingly degrades the model performances on the task. In Figure 4 (left), we observe complete missing of the broken line (top sample) and an image which is almost completely annotated as faulty when multiple broken CCD pixels are present. Nevertheless, no false positives are observed when segmenting fully-nominal images (bottom sample). We conclude that the applied downsampling yields too coarse feature maps to grasp thin structures. As a consequence, images with a prevalence of nominal pixels will be correctly segmented, while isolated thin structures will be absorbed by the majority of nominal pixels around them. Analogously, Figure 4 (right) shows the mostly correct segmentation of the nominal images and a wide spread of the segmentation mask in presence of dust grains (bottom sample). A peculiar bias in the dust segmentation model is partially observed in the top sample of Figure 4 (right), where totally dark areas are misinterpreted as dust grains.

Lastly, data augmentation strategies are introduced for the purpose of scaling and stretching the artifacts, especially broken lines and pixels, as they are one-pixel wide. We adopt *RandomResizedCrop* from the *torchvision* library [28] with the following parameters: $size = 1024$, $scale = (0.125, 0.5)$, $ratio = (1.0, 1.0)$, $interpolation = BILINEAR$ or $interpolation = NEAREST$ for the image and mask respectively. With this data augmentation pipeline it is possible to obtain two to four times the size of the faults, eventually reflecting on the downsampled features. Figure 5 shows that an aggressive scaling and stretching strategy does not significantly improve the segmentation results. Figure 5 shows improved segmentation capabilities of the stretch and rescale approach with respect to varying the $output_stride$ parameter, but it is still not consistently improving performances with respect to the standard model. The model in Figure 5 is capable to profitably detect single broken lines and several broken pixels and dust grains. However, segmented areas are usually larger than the subtended fault, a problem which was already observed in Figure 4. In this case, the model still struggles to separate lines subtended to CCD broken pixels, while it also predicts disjointed dust grains as single artifacts. Finally, nominal pixels segmentation is analogously accurate as in the previous case.

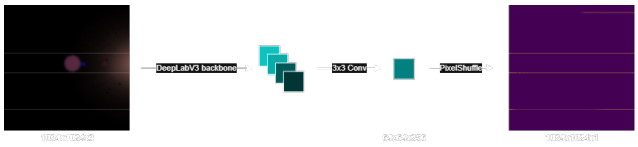

Figure 6. Structure of the augmented DeepLabV3 model. Operations of upsampling are highlighted on the arrows, while the squares are representative of the channel dimension.

Final Model Design The analysis presented earlier in this Section suggests that the segmentation issues are not related to the encoder part of the network and to the data preprocessing stage, because no modification to any module of this block provided relevant improvement with respect to the standard DeepLabV3. Therefore, we focus now on the decoder part of the architecture.

As stated in the DeepLabV3+ work [7], acting on the bilinear upsampling stage of DeepLabV3 may increase segmentation performances on short-scale image features. The use of DeepLabV3+ is deprecated in the present analysis due to the use of skip connections, which imply a higher memory consumption due to the need of storing the features in memory to concatenate. This storage during inference would affect memory-limited devices, i.e., embedded devices targeted in the space domain. Therefore, we introduce a learnable upsampling module similar to DUC presented in [45]. The module entirely substitutes bilinear upsampling of DeepLabV3 with a convolutional layer followed by a *PixelShuffle* layer [36]. *PixelShuffle* redistributes feature matrices to upsample the height and width dimension while reducing the number of channels. The learnable part of the upsampling resides in the convolutional layer, whose function is to match the number of channels needed for *PixelShuffle* to achieve the correct output size. The final architecture is depicted in Figure 6.

The proposed method shows a consistent improvement of results on thin lines and broken pixels , as well as on dust grains segmentation , with respect to the approaches described in the previous paragraphs. Figure 7 highlights better separation of different thin lines with respect to the segmentation masks visualized in Figures 1 to 5. The segmentation of single broken pixels is also improved, with an increased percentage of detected defects. In Figure 8, segmented broken pixels are larger than the corresponding ground truth and exhibit diverse shapes, suggesting that the model is not overfitting the particular shape of the pixels

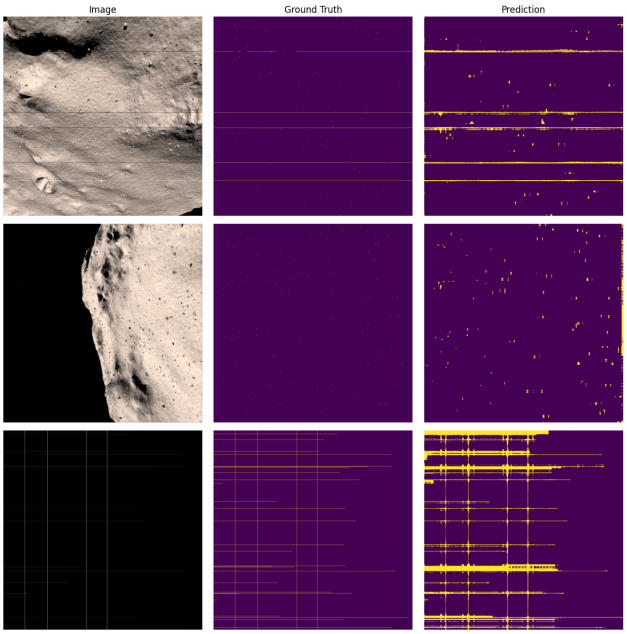

Figure 7. Binary segmentation of Broken Pixels and Lines with the introduced learnable upsampling model.

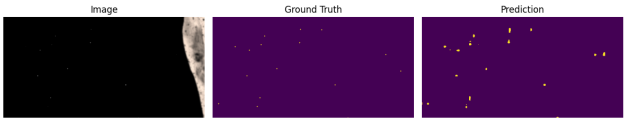

Figure 8. Zoom-in from Figure 7. Single pixels are correctly recognized, with some false positives and false negatives quantified in the metrics.

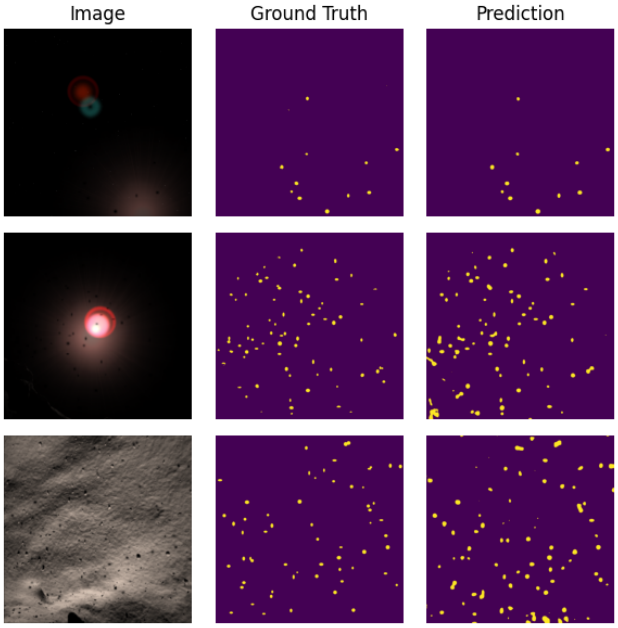

Figure 9. Binary segmentation of Dust on Optics with the introduced learnable upsampling model.

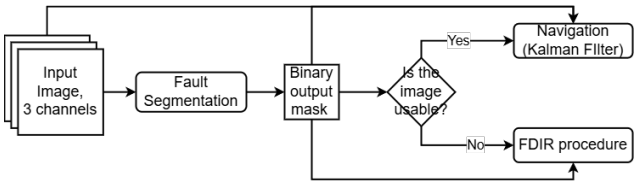

Figure 10. Camera sensor data processing in onboard subsystems

provided in the dataset, as described in [17]. This result suggests that the model may perform well on footprints of subtending pixel defects, originated by e.g., image compression algorithms. Figure 9 generally shows better correspondence of shape and size of predicted and ground truth dust grains with respect to the correspondent previous models in Figures 1 to 5. In the Dust on Optics case, further refinement of the results is achieved employing a combined Binary Cross-entropy and Dice Loss [38], to target both region accuracy and artifacts overlap.

Onboard integration In order to highlight the enhancement brought by the presented segmentation architecture at the spacecraft functional level, we propose a strategy to integrate the neural network output in the onboard processing pipeline. Such pipeline is designed to rely on the fault segmentation results to enhance the system reliability and availability, while making it resilient to the considered faults. As illustrated in the proposed architecture (Figure 10), the segmentation mask identifies faulty pixels and provides valuable information to underweight or exclude corrupted data from the Kalman filter estimation. By isolating these anomalies, the filter maintains its reliance on high-fidelity information, thereby mitigating the risk of divergent or incorrect estimations. Ultimately, a detection logic can be implemented to identify functionally unusable images due to the excessive presence of faulty pixels. This mechanism would trigger an FDIR procedure to restore nominal sensor functionality without compromising the availability of the estimation process. While simple threshold rules on the number of faulty pixels can be implemented to define image *unusability*, the study of more complex approaches is deferred to future work.

The presented onboard processing strategy leverages established architectural frameworks to maximize mission heritage and minimize the design modifications required for AI integration in existing subsystems. Specifically, it builds upon proven spaceflight mechanisms by extending the traditional concept of binary validity flags (conventionally utilized for verifying telemetry according to specific checks)

to the domain of multi-dimensional camera data. By utilizing binary segmentation matrices to mask corrupted pixels, our solution enables a granular data integrity assessment and a fine evaluation of the validity of information contributed by individual pixels, ultimately allowing to preserve the availability of partially degraded sensors while ensuring that downstream AI algorithms operate exclusively on clean data. Ultimately, by delegating the detection and recovery of unusable images to the FDIR subsystem rather than the neural network itself, the proposed approach enhances reliability and minimizes architectural impact. The AI-driven augmentation does not replace, but rather operates in synergy with, FDIR flight-proven procedures.

4. Evaluation and Results

The evaluation of the proposed segmentation model is conducted by analyzing its functional impact on the navigation subsystem, consistent with the methodology outlined in Section 3. This assessment builds upon the integration framework and considerations established in [18], which introduces a procedure for embedding convolutional neural networks within onboard FDIR subsystems for time-series anomaly detection. Central to that framework is a specialized set of case-specific metrics designed to account for both the detection logic and the integration constraints. Although the *System Metrics* defined in the previous study are not directly applicable here — as full hardware integration remains a subject for future work — the underlying principles of the *Detection Metrics* are readily transferable to the domain of image-based fault detection. This adaptation aligns with the recommendations in [18] regarding the extensibility of such metrics to diverse FDIR tasks. Consequently, Table 2 provides a comparative overview of the original metrics and introduces the refined set applied in this study, justifying their selection based on the peculiarities of multi-dimensional camera data.

The interval-based evaluation paradigm proposed in [18] prioritizes the detection of fault-induced artifacts over individual pixels. In the context of computer vision, this approach translates to an artifact-centric evaluation, wherein the primary objective is the identification of contiguous faulty regions rather than per-pixel classification accuracy. This transition is motivated by the specific requirements of the navigation subsystem, where high-fidelity boundary precision is secondary to the reliable exclusion of corrupted information. Standard segmentation benchmarks, i.e., Precision, Recall, and mean Intersection over Union (mIoU), typically yield suboptimal values for the models discussed in Section 3.2 due to the tendency of predicted masks to encompass areas slightly larger than the ground truth artifacts. However, this over-segmentation is not inherently critical for robust navigation; indeed, conservative masking may prove advantageous if onboard image compression or unexpected environmental conditions alter the defect's morphological characteristics relative to the training data, as illustrated in Figure 8. To rigorously quantify this behavior, a comprehensive evaluation framework is established by assessing the overlap between predicted and ground truth regions alongside the magnitude of the size discrepancy, as detailed in the metrics presented in Table 2.

Table 2. Custom metrics for image faults segmentation.

Metric Name		Definition
Per-artifact (PaR)	Recall	$1 - \frac{\text{False Negative Artifacts}}{\text{Total Ground Truth Artifacts}}$
Per-artifact (PaP)	Precision	$1 - \frac{\text{False Positive Artifacts}}{\text{Total Predicted Artifacts}}$
Per-artifact (PamIoU)	mIoU	$\frac{\text{intersection}}{\text{union}}$, True Positives only

The evaluation of artifact prediction performance utilizes established image processing algorithms to delineate specific regions of interest. By employing morphological operators from the *scikit-image* library [42], connected components are extracted from both the annotated masks and model predictions to identify discrete features such as dust grains, broken pixels, and defect lines. These delineated regions are subsequently compared to identify False Positive and False Negative artifacts, which are excluded from the calculation of the PamIoU. The PamIoU is defined as the ratio of intersecting pixels to union pixels between the ground truth and the prediction. Critically, the numerical result of the PamIoU remains invariant whether it is calculated by aggregating individual artifact comparisons or by performing a global pixel-wise summation across the entire frame. To optimize computational efficiency, the latter approach is adopted, thereby bypassing the need for complex data association techniques (i.e., Hungarian algorithm) which would be required to resolve many-to-one mapping scenarios, as illustrated by the clustered dust grains in Figure 9. Consequently, the primary distinction between the PamIoU and the conventional mIoU lies in this selective exclusion of False Positive and False Negative artifacts, focusing the metric exclusively on the segmentation fidelity of correctly identified features.

The quantitative evaluation of the segmentation models, as summarized in Table 3, confirms the performance trends initially expected during the qualitative assessment in Section 3.2. Specifically, results substantiate the diminished efficacy of models utilizing reduced backbone downsampling (Model #2) and aggressive pre-processing (Model #3), while validating the comparable performance levels between the standard and modified Atrous Spatial Pyramid

Dust								
	Loss	ASPP (6,12,18)	$OS = 8$	Stretch and Rescale	Learnable Upsampling	PaP	PaR	PamIoU
Model #0	BCE					0.471	0.468	0.515
Model #1	BCE	✓				0.566	0.532	0.356
Model #2	BCE	✓	✓			0.042	0.711	0.007
Model #3	BCE	✓		✓		0.382	0.514	0.064
Model #4	BCE Dice	✓			✓	0.511	0.526	0.572
Broken Pixels and Lines								
	Loss	ASPP (6,12,18)	$OS = 8$	Stretch and Rescale	Learnable Upsampling	PaP	PaR	PamIoU
Model #0	BCE					0.720	0.611	0.110
Model #1	BCE	✓				0.811	0.425	0.112
Model #2	BCE	✓	✓			0.299	0.540	0.053
Model #3	BCE	✓		✓		0.331	0.748	0.056
Model #4	BCE	✓			✓	0.333	0.942	0.130

Table 3. Custom metrics results for the Dust and Broken Pixels and Lines models.

Pooling (ASPP) architectures (Models #0 and #1, respectively). The proposed final architecture (Model #4) demonstrates superior performance across the majority of metrics and fault categories; however, it exhibits a marginally lower PaP in the broken pixels and lines cases due to the previously discussed tendency toward conservative, larger-area predictions per artifact. These findings indicate that while a robust solution for segmenting complex camera-induced faults is achievable through the Model #4 architecture, further hyperparameter tuning may be necessitated to align the model's precision-recall balance with specific mission-specific requirements.

5. Conclusion and Future Work

In this work, a novel approach for the onboard segmentation of camera faults in satellite applications has been presented, specifically targeting dust on optics and broken pixels or lines. Building upon prior characterizations of camera sensor anomalies, the proposed methodology integrates a custom learnable upsampling module within a DeepLabV3 architecture, significantly enhancing the segmentation of thin and irregular structures compared to alternative configurations employing reduced downsampling rates, aggressive pre-processing, or modified Atrous Spatial Pyramid Pooling (ASPP) modules. The architecture is engineered for compatibility with the stringent memory and computational constraints of spacecraft systems and includes a functional integration framework that defines the interface between the AI-driven detection model and the navigation and FDIR subsystems. Central to this framework is a transition from conventional pixel-wise metrics to an artifact-centric evaluation paradigm, reflecting the operational priority of detecting contiguous fault regions over achieving high-fidelity boundary precision.

Future research shall prioritize the validation of this model through cross-domain verification across diverse mission profiles (e.g., docking and landing) and benchmarking against empirical data from active missions to ensure operational reliability. Furthermore, transitioning from simulated environments to hardware implementation on space-qualified Field-Programmable Gate Arrays (FPGAs) is essential to verify real-time inference capabilities. Ultimately, the integration of this model into a comprehensive navigation pipeline will allow for an assessment of system availability and the definition of operational protocols required to maintain a fail-operational state or trigger autonomous recovery procedures.

References

[1] Gaétan Bahl, Lionel Daniel, Matthieu Moretti, and Florent Lafarge. Low-power neural networks for semantic segmentation of satellite images. In *Proceedings of the IEEE/CVF international conference on computer vision workshops*, pages 0–0, 2019. 2

[2] David S Bayard, Dylan T Conway, Roland Brockers, Jeff H Delaune, Larry H Matthies, Håvard F Grip, Gene B Merewether, Travis L Brown, and Alejandro M San Martin. Vision-based navigation for the nasa mars helicopter. In *AIAA Scitech 2019 Forum*, page 1411, 2019. 1

[3] BJ Bos, DS Nelson, JY Pelgrift, AJ Liounis, D Doelling, CD Norman, RD Olds, CW May, R Witherspoon, E Church, et al. In-flight calibration and performance of the osiris-rex touch and go camera system (tagcams). *Space Science Reviews*, 216:1–52, 2020. 1

[4] Fanghui Chen, Shouliang Li, Jiale Han, Fengyuan Ren, and Zhen Yang. Review of lightweight deep convolutional neural

networks. *Archives of Computational Methods in Engineering*, 31(4):1915–1937, 2024. 2

[5] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017. 2

[6] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017. 3

[7] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818, 2018. 2, 5

[8] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017. 2

[9] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016. 2

[10] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009. 2

[11] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021. 2

[12] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88:303–338, 2010. 2

[13] Frédéric Férésin, Erwann Kervennic, Yves Bobichon, Edgar Lemaire, Nassim Abderrahmane, Gaétan Bahl, Ingrid Grenet, Matthieu Moretti, and Michael Benguigui. In space image processing using ai embedded on system on module: example of ops-sat cloud segmentation. In *2nd European Workshop on On-Board Data Processing*, 2021. 2

[14] European Cooperation for Space Standardization. Telemetry and telecommand packet utilization. Standard ECSS-E-ST-70-41C, 2016. 1

[15] Gianluca Furano, Gabriele Meoni, Aubrey Dunne, David Moloney, Veronique Ferlet-Cavrois, Antonis Tavoularis, Jonathan Byrne, Léonie Buckley, Mihalis Psarakis, Kay-Obbe Voss, et al. Towards the use of artificial intelligence on the edge in space systems: Challenges and opportunities. *IEEE Aerospace and Electronic Systems Magazine*, 35(12): 44–56, 2020. 2

[16] Riccardo Gallon, Fabian Schiemenz, Alisa Krstova, Alessandra Menicucci, and Eberhard Gill. Machine learning-based vs deep learning-based anomaly detection in multivariate time series for spacecraft attitude sensors. In *Proceedings of SPAICE2024: The First Joint European Space Agency/IAA Conference on AI in and for Space*, pages 388–392, 2024. 1

[17] Riccardo Gallon, Fabian Schiemenz, Alessandra Menicucci, and Eberhard Gill. Addressing camera sensors faults in vision-based navigation: Simulation and dataset development. *arXiv preprint arXiv:2507.02602*, 2025. Under Review in Acta Astronautica. 1, 2, 3, 6

[18] Riccardo Gallon, Fabian Schiemenz, Alessandra Menicucci, and Eberhard Gill. Convolutional neural network design and evaluation for real-time multivariate time series fault detection in spacecraft attitude sensors. *Advances in Space Research*, 2025. 1, 7

[19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(9):1904–1916, 2015. 2

[20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 2, 3

[21] Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, et al. Searching for mobilenetv3. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1314–1324, 2019. 2

[22] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. 2

[23] Mate Kisantal, Zbigniew Wojna, Jakub Murawski, Jacek Naruniec, Kyunghyun Cho, et al. Augmentation for small object detection. In *CS & IT Conference Proceedings*. CS & IT Conference Proceedings, 2019. 2

[24] Alexander Liesch, Patrick Suwinski, Bangshang Liu, Valerij Chernykh, and Klaus Janschek. Ai-based lidar/camera data fusion to enable high-resolution 3d surface reconstruction for autonomous asteroid exploration mission. In *AAS/AIAA Astrodynamics Specialist Conference*, 2023. 1

[25] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014, proceedings, part v 13*, pages 740–755. Springer, 2014. 2

[26] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10012–10022, 2021. 2

[27] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015. 2

[28] TorchVision maintainers and contributors. Torchvision: Pytorch's computer vision library. https://github.com/pytorch/vision, 2016. 5

[29] Maurice Martin, Edoardo Caroselli, Jan Olucak, Silvia Busi, Walter Fichter, Bangshang Liu, Alexander Liesch, Patrick Suwinski, Valerij Chernykh, and Klaus Janschek. Pioneering the small bodies frontiers: The key enabling technologies for autonomous precise mobilitys. In *12 th International Conference on Guidance, Navigation & Control Systems (GNC)*, 2023. 1

[30] Sachin Mehta, Mohammad Rastegari, Anat Caspi, Linda Shapiro, and Hannaneh Hajishirzi. Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. In *Proceedings of the european conference on computer vision (ECCV)*, pages 552–568, 2018. 2

[31] Jan Olucak, Sebastian Zähringer, Fabian Schimpf, and Walter Fichter. Sensor based motion planner for small solar system body exploration. In *AIAA SCITECH 2023 Forum*, 2023. 1

[32] Carolyn C Porco, Robert A West, Steven Squyres, Alfred McEwen, Peter Thomas, Carl D Murray, Anthony Delgenio, Andrew P Ingersoll, Torrence V Johnson, Gerhard Neukum, et al. Cassini imaging science: Instrument characteristics and anticipated scientific investigations at saturn. *The Cassini-Huygens Mission: Orbiter Remote Sensing Investigations*, pages 363–497, 2004. 1

[33] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015. 2

[34] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018. 2

[35] Tobias Schwarz, Oleksii Balagurin, Tobias Greiner, Tobias Herbst, Tobias Kaiser, Hakan Kayal, and Andreas Maurer. Early results and inflight experience of the 6U-mission SONATE-2. In *Small Satellites Systems and Services Symposium (4S 2024)*, pages 350–365. SPIE, 2025. 2

[36] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883, 2016. 5

[37] Jason M Soderblom, James F Bell III, Jeffrey R Johnson, Jonathan Joseph, and Michael J Wolff. Mars exploration rover navigation camera in-flight calibration. *Journal of Geophysical Research: Planets*, 113(E6), 2008. 1

[38] Carole H Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*, pages 240–248. Springer, 2017. 6

[39] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015. 2

[40] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019. 2

[41] Mingxing Tan and Quoc Le. Efficientnetv2: Smaller models and faster training. In *International conference on machine learning*, pages 10096–10106. PMLR, 2021. 2

[42] Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014. 7

[43] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 2

[44] Haibo Wang, Wenyong Yu, Jiangbin You, Ruolin Ma, Weilin Wang, and Bo Li. A unified framework for anomaly detection of satellite images based on well-designed features and an artificial neural network. *Remote Sensing*, 13(8):1506, 2021. 2

[45] Panqu Wang, Pengfei Chen, Ye Yuan, Ding Liu, Zehua Huang, Xiaodi Hou, and Garrison Cottrell. Understanding convolution for semantic segmentation. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 1451–1460. Ieee, 2018. 2, 5

[46] Robert West, Benjamin Knowles, Emma Birath, Sebastien Charnoz, Daiana Di Nino, Matthew Hedman, Paul Helfenstein, Alfred McEwen, Jason Perry, Carolyn Porco, et al. In-flight calibration of the cassini imaging science sub-system cameras. *Planetary and Space Science*, 58(11):1475–1488, 2010. 1

[47] Manabu Yamada, Toru Kouyama, Koki Yumoto, Eri Tatsumi, Naofumi Takaki, Yasuhiro Yokota, Tomokatsu Morota, Naoya Sakatani, Masahiko Hayakawa, Moe Matsuoka, et al. Inflight calibration of the optical navigation camera for the extended mission phase of hayabusa2. *Earth, Planets and Space*, 75(1):36, 2023.

[48] RA Yingst, S Bray, K Herkenhoff, M Lemmon, ME Minitti, ME Schmidt, KS Edgett, DM Fey, and LC Kah. Dust cover on curiosity's mars hand lens imager (mahli) calibration target: Implications for deposition and removal mechanisms. *Icarus*, 351:113872, 2020. 1

[49] Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015. 2