SAMPLES DETECTION AND RETRIEVAL FOR A SAMPLE FETCH ROVER

L. M. Mantoani[1], R. Castilla-Arquillo[1], G. J. Paz-Delgado[1], C. J. Pérez-del-Pulgar[1], and M. Azkarate[2]

[1] *Universidad de Málaga, Space Robotics Laboratory, Málaga, Spain; carlosperez@uma.es*
[2] *European Space Agency, ESTEC, Noordwijk, Netherlands; martin.azkarate@esa.int*

ABSTRACT

Future planetary exploration missions are demanding more and more autonomy since these missions are getting more complex. A clear example is the Mars Sample Return mission, where the Sample Fetch Rover needs to collect sample tubes on a remote location, and bring them back to the base station to be launched to Earth. This mission requires to extend the autonomous capabilities onboard. First, the Navigation component needs to be able to detect and locate the sample tubes, and second, the Guidance and Control ones require to place the rover close the sample tubes and move the manipulator to pick them up. These are the main contributions of this paper. The first issue has been solved by the use of Deep Neural Networks, which allow to identify the previously trained sample tubes on images, and the second one has been solved by extending the path planning algorithm within the Guidance component. To demonstrate and validate the proposed methods, two experiments were carried out. A first field test in the Search and Rescue experimental terrain at the University of Malaga, and a second lab test in the Planetary Robotics Lab at the European Space Agency. Both experiments were carried out using the ExoMars Testing Rover owned by the last institution.

Key words: sample retrieval; Sample Fetch Rover; planetary exploration.

1. INTRODUCTION

Planetary exploration is requiring more and more autonomy for rover to increase scientific return and be able to perform even more complex missions. The main responsible of it is the Guidance, Navigation and Control (GNC) architecture, which is key, since the remote teleoperation from Earth entails several difficulties that can be effectively overcome with autonomous onboard decision making [1]. As a part of this architecture, the Guidance component is designed to plan a path for the planetary exploration vehicle, allowing the rover to navigate safely, avoiding any hazard that could be found along its way. The Navigation component is in charge of detecting these obstacles and locating the rover, which would ensure it is following the right path. To ensure the generated path is correctly followed, the Control component is responsible of generating actuator commands according to the rover location. The importance of the autonomous decision making depends on the objectives of the exploration mission. In the case of the future Sample Fetch Rover (SFR) that belongs to the Mars Sample Return (MSR) mission, the objective is to collect several soil sample tubes left by the M2020 Perseverance rover and bringing them back to the lander in 150 sols [2]. In order to increase the overall navigation speed of the system, it is necessary to perform the sample retrieval operation in the most efficient, hereby efficient, way. Thus, the GNC architecture needs to be extended to solve two issues: first, an efficient GNC architecture that allows the rover to reach the sample locations faster than before, and second, a path and motion planner to retrieve the sample tubes and store them. As regards the first issue, authors proposed an efficient GNC architecture that would be suitable to accomplish the stated objectives [3]. The design and development of the second issue is the aim of this contribution.

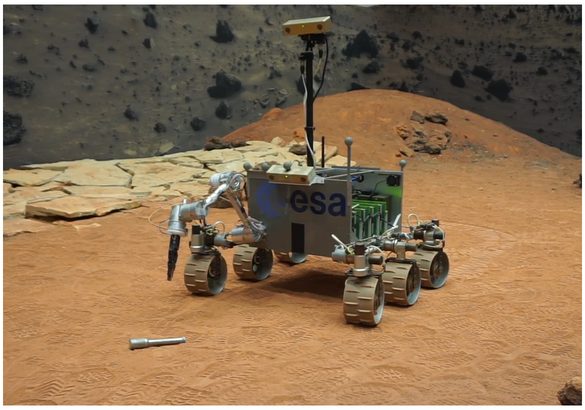

Figure 1: ExoTeR rover during the SFR field test campaign at ESA-ESTEC.

To accomplish the MSR objectives, the Navigation subsystem needs to be extended to perform automatic detection and localization of the sample tubes. For this purpose, the use of deep neural networks (DNNs) is one the best techniques nowadays. DNNs perform some tasks better than humans, as it is the case of image classification [4]. However, the implementation of DNNs in space-related projects can pose some challenges. A high volume of images is needed to train DNNs to accurately recognise the sample tube. This problem can be solved with the creation of a synthetic dataset with inmersive 3D simulations of martian scenarios. The knowledge obtained while training from this generic dataset is later transferred to the DNN to be used in more realistic environments, such as the field tests. Another challenge is related to the high computational demands of DNNs. Therefore, the trained network is adapted and implemented into a hardware-accelerated device, with energy efficiency in mind [5]. For the pose estimation stage, traditional computer vision techniques are developed and integrated.

Once the sample is detected and its pose estimated, the rover needs to reach the samples location and retrieve it. Due to the high accuracy requirement for a successful retrieval, the sample is repeatedly detected during the execution, to feedback more accurately its pose. Besides, the continuous sample identification makes it possible to ignore, for instance, the rover localization errors, which would lead, otherwise, to a mission failure. This way, the Guidance subsystem can plan a path to reach the samples vicinity, and correct it with every new detection if necessary. Besides, Guidance needs to handle the mobile manipulation problem, placing the rover base in the best pose for the retrieval operation and moving the robotic arm in a correct manner to pick up the sample tube.

The proposed sample retrieval methodology has been validated by means of simulation tests in a representative virtual environment built on Unreal Engine, a field test campaign in the Search and Rescue Experimental Terrain at the University of Málaga, and a lab test in the Planetary Robotics Laboratory at the European Space Agency. The rover testbed ExoTeR (Exomars Testing Rover) [6] has been used. As shown in figure 1, it is equipped with a 5 Degrees of Freedom (DoF) manipulator, a gripper and two stereo cameras, NavCam and LocCam. These cameras are used for localization, mapping and sample detection. The performed field tests showcased an analogue SFR mission, where ExoTeR was able to autonomously detect and locate a sample tube, reaching it and picking it up with the manipulator. An already ESA developed visual odometry algorithm was used for the rover localization, avoiding the use of GNSS devices, which increases the representativeness of the tests. The obtained results show that the proposed autonomous sample detection and motion planning architecture was able to solve efficiently the sample retrieval stage of a SFR-like mission.

2. SAMPLE LOCALIZATION

The sample localization procedure has been divided into two stages: sample identification and pose estimation (position and orientation). In the sample identification stage, images from the four sensors of both LocCam and NavCam were employed to search for the sample tube. A DNN was implemented to localise the sample tube in the scene, thus obtaining a bounding box of its position on the image. On the other hand, pose estimation was performed using the LocCam images once the rover is prepared to pick up the sample tube. As the rover has stereo cameras as its disposal, disparity maps were used to transform the image coordinates provided by both algorithms to real world 3D coordinates, which were fed to the GNC algorithm.

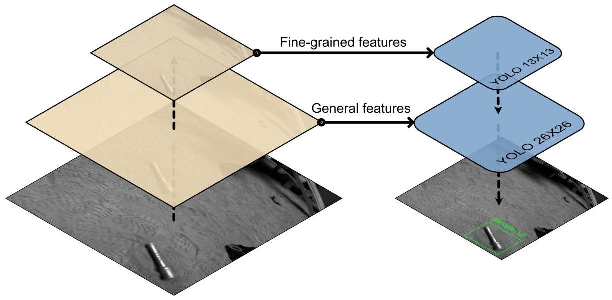

Figure 2: Pyramid Feature Extractor. On the left, feature maps of different resolution are extracted. On the right, bounding boxes are calculated for each feature map and their outputs are fused into a final bounding box.

YOLOv3-tiny architecture, a smaller version of YOLOv3 [7], was the DNN implemented for the sample identification. It is based on the Darknet framework, which, in turn, is heavily influenced by the Pyramid Feature Extraction (PFE) procedure [8], as seen in Fig. 2. It extracts the images characteristics through consecutive convolutional layers of decreasing dimensions, creating two feature maps of different level of detail. Additionally, as an output of its outer layers (decoding head), boundings boxes are precomputed and the final ones are chosen depending on the predicted probabilities. This type of feature decoding, along a reduced number of layers, contributes to a faster and less dense network. This makes easier its adaptation in hardware-accelerating devices, enhancing the detection times of resource-constrained CPUs.

Generally, this type of DNNs needs large datasets of images for an efficient training and a high ratio of successful predictions. Nonetheless, obtaining a large enough dataset of the sample tube in real conditions is rather complex. Transfer learning techniques are proposed to solve this issue [9]. A dataset was produced with a high volume of synthetic images. In them, the sample tube was represented in a wide range of scenarios and illumination conditions. For this task, a photorealistic simulator based on Unreal Engine was used. In addition, a second dataset was generated with real images of the sample tube in a terrain similar to that of the martian surface. Then, the YOLOv3-tiny network was first trained with the synthetic dataset to initialise its weights. The first layers of the pretrained network were later transferred to a second network trained with real images, keeping frozen the previous trained layers. Thanks to this procedure, the network is able to learn the sample tube general features of the synthetic dataset, making the network more robust to condition changes and increasing the ratio of successful

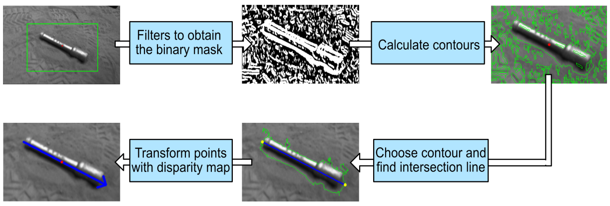

Figure 3: Diagram of the algorithm used to compute the sample tube pose. Its input is the bounding box provided by the YOLOv3-tiny network.

detections.

The DNN was adapted to a hardware accelerator device. Specifically, it was implemented into a Google's Coral Tensor Processing Unit (TPU) as it has good performance in terms of power consumption per inferenced image [10]. A quantization of the YOLOv3-tiny tensor data was performed from a 32-bit to a 8-bit fixed-point representation. Additionally, some layers and operations were adapted to similar ones supported by the TPU, for example, leaky Rectified Linear Unit (ReLU) were changed by standard ReLU layers. Although these changes reduced the network accuracy due to rounding errors, they were compensated by the transfer learning procedure.

Computer Vision (CV) techniques were employed for the the pose estimation algorithm. A diagram of all the followed operations are depicted in Fig. 3. Starting on top left, the bounding box of the detected sample tube is introduced as the algorithm input. The sample tube binary mask is therefore obtained via the combination of a Sobel filter to detect its edges and a adaptive mean threshold filter. Consecutively, all the images contours and their centroids are generated. Among them, the contour chosen as that of the sample follows two conditions: its centroid is the nearest to the one detected by the DNN and it is enclosed by the contour. In addition, a line is calculated which fits all the contour points using the least-squares algorithm. This is followed by the estimation of the two points of intersection of the contour with the approximated line. Finally, these points are translated to real world coordinates using the LocCam disparity maps and the pinhole camera model.

3. SAMPLE RETRIEVAL PLANNING AND CONTROL

The sample retrieval planning and control subsystem is divided into two different components, the Path Planner [11] and the Trajectory Control [12]. The Path Planner is in charge of the generation of the optimal path to reach a goal position based on a global map (from drone or orbital imagery), while the Trajectory Control uses the conservative pure pursuit method, to guarantee that the rover will follow the generated path within the limits of a predefined corridor.

The functioning of the Path Planner is the following. Firstly it is necessary to provide a map of the scenario in form of a Digital Elevation Map (DEM), which is processed to differentiate obstacles and safe areas. Hence, a Cost Map is generated based on the traversable areas and slopes encountered in the DEM, where all non-safe areas are dilated in order to ensure the rover will not traverse nor approach them. Clearly, the highest costs in the Cost Map correspond to obstacles. After obtaining the Cost Map from the DEM provided, a path is computed: taking into account the Cost Map, the rover initial position and the first estimation of the sample position, a path is generated for the rover to follow. The trajectory is planned by means of the Fast Marching Method (FMM), which, given a Cost Map of the scenario, always finds a smooth, continuous and optimal trajectory. Finally, when the complete path is obtained, it is shortened by a distance that depends on the arm kinematic characteristics, deleting the last waypoints to ensure that the sample is reachable for the arm.

Each time a new path is computed, it is sent to the Trajectory Control component, which continuously generates, in function of the current rover pose, the motion commands, i.e. translational and rotational speeds, to let the rover follow the generated path. The Trajectory Control library guarantees the rover remains within the limits of a safety corridor. All the speed commands generated are sent to the Locomotion Control component, which translates them into wheel speeds commands. It is important to remark that the last waypoint heading must leave the rover facing the sample to ensure that the arm is able to retrieve it. Hereby, once this waypoint is reached, the Trajectory Control may be forced to command a spot turn, to get the rover to the desired heading.

The planning and control algorithms need to receive the goal position from the sample identification subsystem, which is integrated as follows. SFR is expected to be equipped with two different cameras, a Navigation Camera (NavCam), placed on the top of a mast, and a Localizacion Camera (LocCam), placed on the top of the rover body. Making most of the cameras, continuous sample identifications are requested while the rover is following the planned path, updating the sample position to a most accurate one. If the sample identification subsystem manages to recognise the sample, a new final sample position is provided, from which a new path is calculated. Otherwise, the rover will continue following the previous path, periodically requesting new detections until the sample is recognised or the final waypoint is reached. Given its height, the NavCam has a wider field of view. For this reason, the first sample identifications are performed by the NavCam, until the rover is close enough to request a LocCam detection. Generally, at least three successful sample identifications are expected.

Finally, when the rover reaches the last waypoint of the path, it requests a highly accurate sample recognition, before starting the retrieval with the robotic arm. The inverse kinematic model of the manipulator is used, then, to obtain a motion plan for the gripper. Firstly it reaches a point just above the sample, and then smoothly descends to perform the picking operation.

4. EXPERIMENTAL RESULTS

In order to validate and demonstrate the proposed autonomous sample detection and retrieval architecture for a real Sample Fetch Rover mission, several tests were performed with the Exomars Testing Rover (ExoTeR). As aforementioned, ExoTeR is a triple-bogie, six-wheeled, full-Ackermann rover, which replicates the locomotion subsystem of the Exomars Rosalind Franklin Rover, including a LocCam, a NavCam and a 5 DoF manipulator, as can be observed in Figure 1. Two different test campaigns were carried out. On the one hand, the simulation tests within a tailored martian simulation environment, which uses Vortex Studio as a physical engine and Unreal Engine 4 as the visualisation engine. On the other hand, the experimental validation of the algorithms in real environments with, first, field tests in a representative outdoor environment in the Search and Rescue experimental terrain located at the University of Málaga, and, second, laboratory tests in the analogue martian testbed located in the Planetary Robotics Laboratory at ESA-ESTEC.

Starting with the simulation tests, Unreal Engine 4 allowed to test computer vision algorithms, as it was the case for training the DNN as aforementioned. Second, the powerful physics engine provided by Vortex Studio generates highly reliable motions and interactions of the robotic system with the environment. As can be observed in Figure 4, a representative scenario of a SFR mission was created, consisting of a 9x9 m martian analogue terrain with features like rocks, dust or slopes, including a sample tube and the rover ExoTeR equipped with a 6 DoF manipulator and a two-fingered gripper. Thus, several sample detection and retrieval tests were performed within the simulation environment[1]. Following the software-in-the-loop concept, the simulation tests validated the proposed algorithms in two main senses: the sample detection subsystem properly identified and located the sample w.r.t. the rover pose, using the NavCam for long distance detections and the LocCam for close range ones; and the autonomous navigation subsystem was capable of properly generate and follow a trajectory to reach the identified sample, even if the initial commanded sample pose included a substantial error w.r.t. the real one, and without stopping the rover during the sample detection operations.

Regarding the experimental validation of the algorithms in real environments, two different scenarios were used, to verify if the sample identification subsystem is susceptible to their differences and to test how the whole system behaves in different situations. For this experimental tests, an electrically powered two-fingered gripper was attached to the end effector of the ExoTeR MA5-E 5 DoF manipulator for the retrieval operations.

[1] https://youtu.be/sSDGohU7X7w

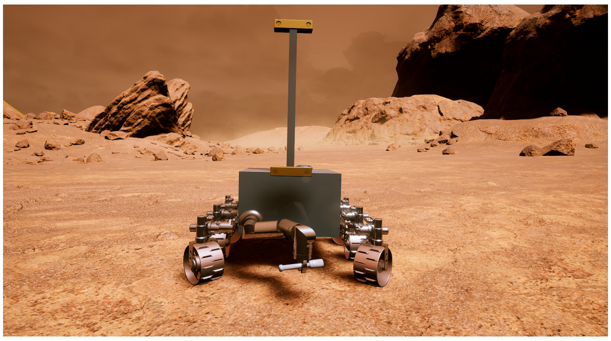

Figure 4: Mars analogue simulation environment on Unreal Engine 4, including the model of ExoTeR and the sample tube.

Note that the 5 DoF manipulator is not able to reach any commanded Cartesian position and orientation, thus, a ground-perpendicular gripper orientation was fixed to maximize the system safety. Additionally, to ensure that the tests were as representative as possible of a real SFR mission, a visual odometry algorithm [13] was used for the rover localization, with an average of 7.5% drift in position and less than 2° orientation. Nevertheless, as will be observed later, thanks to the proposed autonomous sample detection and path planning pipeline, this rover localization error did not influence the sample retrieval, since each time the sample tube was identified, the error was corrected.

The outdoor field tests were carried out at the Search and Rescue experimental terrain located close to the Industrial Engineering Faculty at the University of Malaga. This is a huge terrain which is representative w.r.t. the martian surface conditions, including features like steep slopes, rocks or different types of soil (rough, rocky, sandy...). For performance analysis purposes, the ground truth pose of the rover was obtained by means of a differential GPS with RTK corrections using a fixed antenna. For these tests, a new set of the sample tube images on the experimental terrain were included to the sample identification DNN, on top of the pretrained synthetic model, to consider the particular characteristics of the scenario. It is important to note that the light conditions affected noticeably the behaviour of the sample identification subsystem, since the intense incidental light were occasionally saturating the cameras, as can be observed in the videos recorded by the LocCam and NavCam during the tests[2]. This was partly mitigated by the pretrained model, which contained synthetic images from different day times and different orientations w.r.t. the Sun position, in order to consider a wider variety of light conditions. As a result, the sample identification subsystem was able to recurrently and reliably locate the sample with the NavCam, from approximately 5 m far. Besides, the subsequent LocCam sample identifications eliminated the localization error due to visual odometry, thus, allowing the rover to precisely retrieve the sample tube with the gripper.

[2] https://youtu.be/6nDSPn1uV7E

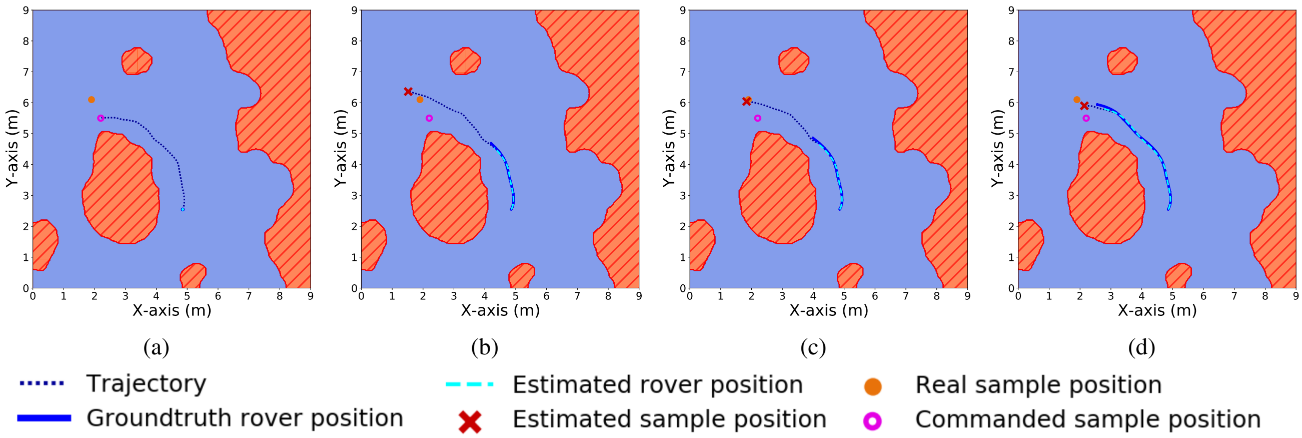

Figure 5: Evolution of one of the Planetary Robotics Laboratory tests including the initial state (a), the NavCam sample identification (b), one of the two consecutive LocCam sample identifications (c) and the final state (d) before the sample tube retrieval. Although the commanded sample position is slightly different from the real one, as long as the rover starts following the generated trajectory the estimated sample position keeps improving after each successful sample identification. Before the retrieval (d), the estimated sample position w.r.t. the rover pose is very accurate, although there is an error in the rover pose w.r.t. its ground truth pose, due to the drifted rover localization.

Lastly, a final laboratory tests campaign was carried out at the Martian Analogue Testbed of the Planetary Robotics Laboratory, at ESA-ESTEC, to verify the behaviour of the already trained neural network and also confirm the robustness of the proposed sample retrieval planning pipeline. The Martian Analogue Testbed is a 9x9 m terrain that replicates very accurately the martian soil conditions and features, as can be observed in Figure 1. Inside the PRL, a set of Vicon markers were placed on ExoTeR to obtain its ground-truth pose during the tests. The evolution of one of the lab tests is depicted in Figure 5. Initially, Figure 5a, the rover is commanded to reach and retrieve a sample (purple o), including a certain error from the real sample position (orange circle), starting from its initial pose (blue dot). Thus, the rover generated an initial trajectory (dark blue dashed line) to reach the sample tube vicinity. Nevertheless, once the rover had covered a few meters of the trajectory, the NavCam identified the sample tube, Figure 5b, which was located quite far from what was initially expected, and a new trajectory was generated. Then, a first LocCam identification, Figure 5c, increased the accuracy of the sample pose estimation. Lastly, a second LocCam identification allowed the rover to finally locate itself in an appropriate pose for the subsequent sample retrieval operation. Therefore, the final state, Figure 5d, shows how the relative pose of the sample w.r.t. the rover was highly accurate in comparison with the real ones, even though the rover had a certain localization error due to the use of visual odometry. A picture of the final sample pose estimation on both LocCam sensors is depicted on Fig.6.

5. CONCLUSIONS

The objective of this work was the validation and demonstration of an extended GNC architecture for a Sample Fetch Rover. This extension consisted on the use of the rover cameras to detect the sample tubes for Navigation, and the modification of the Guidance component to be able to reach the sample location and move the manipulator to pick it up.

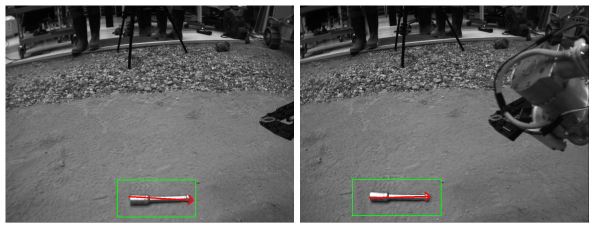

Figure 6: Detection box (green rectangle) and pose estimation (red arrow) of the laboratory test left and right LocCam sensors.

For this purpose, three experiments were carried out. The first one in an inmersive simulation environment, in which obtained images were used to train a DNN to locate the sample tubes. Moreover, it was used to test the proposed GNC architecture using the software-in-the-loop concept. The second experiment was carried out in an experimental terrain using a real rover, where we demonstrated the proposed architecture can be used on a real rover testbed. Finally, another experiment was carried out in a Martian analogue lab facility. The aim of this experiment was to validate the proposed architecture in a more Mars-like terrain, and even to be able to carry out future works. We can conclude the proposed architecture is suitable to be investigated for further planetary exploration space missions such as the SFR. As part of these future works, we propose the optimisation of the rover motion to pick up the sample tubes using novel kinodynamic planning adapted to this use case.

ACKNOWLEDGMENTS

This work was partially supported by the European Space Agency under contract number 4000118072/16/NL/LvH/gp - 3.

REFERENCES

[1] Dan Lester and Harley Thronson. Human space exploration and human spaceflight: Latency and the cognitive scale of the universe. *Space Policy*, 27(2):89–93, 5 2011.

[2] Brian K. Muirhead and Ashley Karp. Mars Sample Return Lander Mission Concepts. In *IEEE Aerospace Conference Proceedings*, volume 2019-March. IEEE Computer Society, 3 2019.

[3] Martin Azkarate, Levin Gerdes, Luc Joudrier, and Carlos J. Perez-del Pulgar. A GNC Architecture for Planetary Rovers with Autonomous Navigation. pages 3003–3009. Institute of Electrical and Electronics Engineers (IEEE), 9 2020.

[4] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 12 2021.

[5] Jiasi Chen and Xukan Ran. Deep learning with edge computing: A review. *Proceedings of the IEEE*, 2019.

[6] Martin Azkarate, Levin Gerdes, Tim Wiese, Martin Zwick, Marco Pagnamenta, Javier Hidalgo Carrio, Pantelis Poulakis, and Carlos Perez-del Pulgar. Design, testing, and evolution of mars rover testbeds: European space agency planetary exploration. *IEEE Robotics & Automation Magazine*, 2022.

[7] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[8] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.

[9] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

[10] Leandro Ariel Libutti, Francisco D Igual, Luis Pinuel, Laura De Giusti, and Marcelo Naiouf. Benchmarking performance and power of usb accelerators for inference with mlperf. In *2nd Workshop on Accelerated Machine Learning (AccML), Valencia, Spain*, volume 31, 2020.

[11] J. Ricardo Sánchez-Ibánez, Carlos J. Pérez-del Pulgar, Martin Azkarate, Levin Gerdes, and Alfonso García-Cerezo. Dynamic path planning for reconfigurable rovers using a multi-layered grid. *Engineering Applications of Artificial Intelligence*, 86:32–42, 11 2019.

[12] Jan Filip, Martin Azkarate, and Gianfranco Visentin. Trajectory control for autonomous planetary rovers. *ASTRA 2017 - 14th ESA Workshop on Advanced Space Technologies for Robotics and Automation*, 2017.

[13] Andreas Geiger, Julius Ziegler, and Christoph Stiller. StereoScan: Dense 3d reconstruction in realtime. *IEEE Intelligent Vehicles Symposium, Proceedings*, pages 963–968, 2011.