(Preprint) AAS 23-422

AI-BASED LIDAR / CAMERA DATA FUSION TO ENABLE HIGH-RESOLUTION 3D SURFACE RECONSTRUCTION FOR AUTONOMOUS ASTEROID EXPLORATION MISSION

Alexander Liesch, Patrick Suwinski, Bangshang Liu, Valerij Chernykh[*], and Klaus Janschek [†]

For future science and exploration operations close to surfaces of Small Solar System Bodies (SSSBs) a high-resolution 3D terrain model is needed for Guidance, Navigation, and Control (GNC) tasks. The relatively low resolution of available flash-LiDAR devices mainly limits the resolution of the detected SSSB's 3D terrain. This paper proposes a method to perform a deep learning-based fusion of low-resolution depth images and high-resolution monocular grayscale 2D camera images to overcome this limitation and increase the resolution of the 3D terrain data acquired by flash-LiDAR. We use a Generative Adversarial Network (GAN) based architecture to process the 3D terrain data of the irregular and unstructured surface of asteroids or comets. A synthetic dataset of $10,000$ samples based on comet 67P/Churyumov-Gerasimenko was generated by using a high-fidelity rendering software for training and validation. Our method is suitable for wide-angle lens applications and shows robustness to varying illumination conditions. A resolution increase by a factor of 8×8 was achieved.

INTRODUCTION

Problem description

Robotic missions to asteroids and comets will be one of the most important areas of solar system exploration in the near future. Current missions to the close vicinity of small celestial bodies require plenty of human monitoring and processing to ensure a safe flight during the space exploration mission. Autonomous operation for exploration of Small Solar System Bodies (SSSBs) with long distances to the Earth is crucial because of particularly impossible real-time communication between the Earth and the exploring spacecraft. Recent space missions have been incorporating increasingly advanced technologies to overcome the limitations of human involvement. Hayabusa-2,[1] launched in 2014, utilized upgraded navigation tools and advanced characterization techniques such as radiometric tracking and autonomous descent. OSIRIS-REx,[2] launched in 2016, took it a step further by incorporating vision-based navigation for close-range operations, advanced exposure techniques, and landmark tracking in addition to radiometric tracking. Most recently, DART[3] launched successfully in 2021 and achieved kinetic impact deflection using a fully autonomous navigation system and avionics.

[*]Research Associate, Institute of Automation, Faculty of Electrical and Computer Engineering, TUD Dresden University of Technology, 01187 Dresden, Germany

[†]Professor, Institute of Automation, Faculty of Electrical and Computer Engineering, TUD Dresden University of Technology, 01187 Dresden, Germany

The recently finished Astrone[4] project funded by the German Space Agency (Deutsches Zentrum für Luft- und Raumfahrt (DLR)) proposed a novel concept of a low-altitude hovering vehicle, which enables advanced surface mobility for science and exploration operations on surfaces of SSSBs like the comet 67P/Churyumov–Gerasimenko. A flash-LiDAR-aided inertial navigation system has been developed and successfully tested in a simulated SSSB environment.[5] As one result of the tests, we identified opportunities to further improve the performance of the entire Guidance, Navigation, and Control (GNC) system by increasing the resolution of the 3D terrain reconstruction. Mainly, the relatively low resolution of available flash-LiDAR devices limits the resolution of the detected 3D terrain. The information of high-resolution monocular grayscale 2D camera images can be used to increase the 3D terrain resolution.

Related Work

Processing the surface of an asteroid or comet poses a significant challenge due to its irregular shape. While a Shape-from-Shading (SfS) approach has been utilized by Gaskell et al. (2008)[6] and Al Asad et al. (2021)[7] in exploration missions to SSSBs, it is necessary to take multiple images of the same scene under different illumination conditions to minimize slope determination ambiguities and albedo variation effects. However, in the Astrone exploration concept, single monocular grayscale 2D images were captured using wide-angle optics with the camera points off-nadir, resulting in strong illumination variations, obscuring, and shadows within the $65^\circ \times 65^\circ$ Field-of-View (FoV). It is not possible to take multiple images of the same scene under different illumination conditions. Considering all these factors, the SfS approach is unsuitable for further investigations following Astrone's exploration concept.

In recent years, different machine learning approaches, such as artificial neural networks, achieved game-changing results in computer vision tasks such as classification, segmentation, super-resolution, and 3D reconstruction. Especially for 3D reconstruction, monocular depth estimation, depth super-resolution, or depth completion have become active research fields with different, but mainly for terrestrial applications. Generally, these methods can be classified by their training mechanism, with supervised, unsupervised (and semi-supervised), and reinforcement learning being the most common. However, high-resolution labeled image-depth datasets for supervised learning approaches are rare, especially for our application.

Further, the existing approaches can be categorized by their input data. Given only an input depth image, unguided methods aim to increase the input's resolution directly. On the other hand, guided strategies use additional data like an image. A guided strategy can be considered for our application since we use an additional monocular grayscale 2D image. The type of data fusion used in these approaches makes it possible to categorize even further. Deep learning-based data fusion mainly consists of early, late, and hybrid fusion.[8–12] We focus on works with a comparable application area and supervised learning methods, which are more relevant to our proposed work. Deep learning-based depth estimation, super-resolution, or depth completion methods were effectively applied to planetary orbital datasets. The Mars3Dnet[13] used a Convolutional Neural Network (CNN)-based structure to predict a high-resolution Digital Elevation Model (DEM) from one monocular grayscale 2D image and subsequentially fused this with a low-resolution DEM. Synthetic and real data, including context camera (CTX) images of the Martian surface, were used for training. Another CNN-based approach was given by Chen et al. (2022),[14] where a high-resolution image and a low-resolution DEM of the lunar surface were used to predict a high-resolution DEM. Then SfS was applied to improve details.

GANs were introduced by Goodfellow et al. (2014)[15] and have recently shown their effectiveness in this field. GANs operate by training a generative network while training a discriminator to distinguish between prediction and ground truth in parallel. After training, only the generator part of the GAN is used to predict the output. This architecture has demonstrated its effectiveness in tasks like monocular depth estimation in recent works.[16–18] Like the pure CNN-based methods, these networks operated on orbital Martian or Lunar datasets to predict high-resolution DEMs. All mentioned approaches were primarily utilized for long-range observations using narrow-angle optics and under constant illumination conditions.

AI-BASED LIDAR/CAMERA DATA FUSION

Network Structure

Our approach employs an advanced GAN architecture to create high-resolution depth images from low-resolution depth images obtained through LiDAR and a monocular grayscale 2D camera. A depth image is a two-dimensional representation of a scene, where each pixel encodes the distance or depth information from the sensor to the corresponding surface point in the scene. To enhance the low-resolution depth image, we use the corresponding high-resolution grayscale 2D image as a guide to add missing high-frequency components for the upscaling of low-resolution depth images.

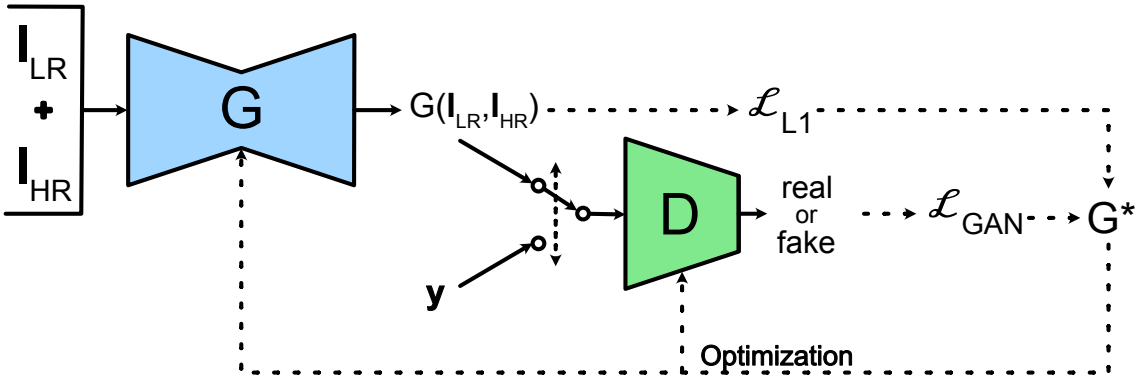


Figure 1. Visualization of the GAN architecture during the training process with the low-resolution depth image $\mathbf{I}_{LR}$, the high-resolution grayscale 2D camera image $\mathbf{I}_{HR}$ and the Loss-functions $\mathcal{L}_{GAN}$, $\mathcal{L}_{L1}$

Our GAN consists of a Generator Network (GN) based on an advanced U-Net[19] architecture and a Discriminator Network (DN). We applied the U-Net3+[20] as an advanced version of the original U-Net using full-scale skip connections as the GN. The input for GN is different from the typical random vector used in GANs. It includes the low-resolution depth image $\mathbf{I}_{LR}$ and the high-resolution 2D grayscale image $\mathbf{I}_{HR}$. The GN's input has the dimension $H \times W \times 2$, where H is the height, and W is the width of the input data. Because of the different dimensions of the low-resolution depth image and high-resolution grayscale 2D image, we scaled the low-resolution depth image to the grayscale image resolution while considering a sparse high-resolution depth image lacking information. By concatenating the grayscale image data and the sparse depth image, we get a concatenated image with a dimension of $H \times W \times 2$ for the GN's input. Please refer to the system processing section for more details on this preprocessing step. Mainly, we trained the GN (UNet3+) to reconstruct the rescaled low-resolution depth image's lacking information.

Since asteroids or comets have irregular surfaces with numerous high-frequency structures, we utilize the CNN-based PatchGAN[21] as the DN. The PatchGAN uses local image patches of the high-resolution depth images generated by the GN or the ground truth data. The DN classifies each 70×70 patch as real (ground truth) or fake (GN's output). Structures at the scale of these patches are penalized. In this case, the DN has the task of restricting the model's high-frequency structures. Further, this approach enables possibilities to judge the validity of the GN's output concerning local regions. As recommended by the authors of PatchGAN,[21] all activation functions of the PatchGAN are Leaky ReLUs with a slope of 0.2. Fig. 1 shows the final GAN structure, including the GN, DN, and their relations.

Objective function

Following the GAN framework, visualized in Fig. 1, the objective function can be expressed by Eq. 1, where the GN's function G tries to minimize the objective against an adversarial, the DN's function D, which tries to maximize it. Here $\mathbf{I}_{LR}$ represents the low-resolution depth image, $\mathbf{I}_{HR}$ the high-resolution grayscale 2D camera image, and $\mathbf{y}$ the ground truth. The mean or expectation $\mathbb{E}$ over all possible input and ground truth data samples of the training dataset is considered.

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{\mathbf{y}}[\log_e D(\mathbf{y})] + \mathbb{E}_{\mathbf{I}_{LR},\mathbf{I}_{HR}}[\log_e(1 - D(G(\mathbf{I}_{LR}, \mathbf{I}_{HR}))] \tag{1}$$

The generator should not only fool the discriminator but also be close the ground truth to produce realistic outputs. For that, Isola et al. (2018)[21] suggested using the $L1$ distance, which encourages less blurring than the $L2$ distance. $L1$ will force low-frequency correctness. Although this will fail to capture details and high frequencies, the discriminator network will take care of this:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{\mathbf{y},\mathbf{I}_{LR},\mathbf{I}_{HR}}[\mathbf{y} - G(\mathbf{I}_{LR}, \mathbf{I}_{HR})] \tag{2}$$

Finally, we get the following objective function G^* for our GN, where the parameter λ limits the influence of $\mathcal{L}_{L1}$:

$$G^* = \arg\min_G \max_D \mathcal{L}_{GAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \tag{3}$$

Training dataset and Details

Open source available labeled datasets for supervised training of CNNs with realistic and usable 2D image and depth image data from asteroids or comets are rare. Thus, in the context of the Astrone project, we applied a high-fidelity modeling software CamSim[22] to generate the grayscale 2D images and the corresponding depth images with a resolution of 1024×1024 from simulated asteroid surfaces. To create a dataset comparable to the current Astrone KI project conditions, we used CamSim to generate a Digital Elevation Map (DEM) inspired by the comet 67P/Churyumov-Gerasimenko's surface, with a resolution of $0.05\,\mathrm{m/pixel}$. We rendered grayscale 2D images and the corresponding depth images with the open-source 3D computer graphics software Blender.[23] The grayscale 2D images were rendered with Lambertian scattering and constant albedo for all surface elements. In rendering, the virtual camera has a FoV of $65° \times 65°$ and imitates the spacecraft's movement during operation at a height between $7\,\mathrm{m}$ to $7\,\mathrm{m}$ over the surface. The sun is initially positioned directly behind the camera.

Further, we generated images with different sun positions. The sun azimuth between $\pm 80^\circ$ within 10° steps created varying amounts of shadows on the surface visible in the additionally generated grayscale 2D images. To simulate the flash-LiDAR data, the corresponding depth images were gathered too. It was assumed that the virtual monocular 2D camera and the virtual flash-LiDAR sensor were perfectly calibrated. Subsequently, the data generation process used the high-resolution depth images as ground truth data to create low-resolution depth images with 128×128 pixels, corresponding to the flash-LiDAR sensor resolution in the Astrone exploration concept. The down-scaling extracted every eighth pixel of the original ground truth data. The total dataset consists of $10,000$ individual samples (independent positions with different sun azimuths), 70% of which were used for training. To form a usable training data set, additional steps according to the following workflow have been implemented:

1. *Sparse depth image creation*: To match the resolution of the input grayscale 1024×1024 2D images, this step created sparse 1024×1024 depth images from the low-resolution 128×128 depth images where missing values are set to 0.
2. *Tiling and Rescaling*: Due to limitations of the used hardware during training, the 2D images and sparse depth image with the dimension 1024×1024 had to be separated into 256×256 tiles to feed it into the GN. Subsequently, a min/max normalization following Eq. 4 rescaled the sparse depth, ground truth depth, and grayscale 2D image tiles to relative floating-point values $f \in [0, 1]$.
3. *Concatenation and Tiling*: The grayscale and corresponding sparse depth were concatenated to a $256 \times 256 \times 2$ tensor, representing the GN's input.

$$\mathbf{x}_{scaled} = \frac{\mathbf{x}_{unscaled} - \min(\mathbf{x}_{unscaled})}{\max(\mathbf{x}_{unscaled}) - \min(\mathbf{x}_{unscaled})} \tag{4}$$

Since the DN operates on the GN's output and ground truth data, its input has a dimension of $256 \times 256 \times 1$. Because of these steps, the number of training samples for the network increases by factor 16.

The GAN was trained for 10 epochs with $121,000$ iterations each. The batch size was set to 10 and an Adam optimizer with a learning rate of $l = 2 \cdot 10^{-4}$, $\beta_1 = 0.9$, and $\beta_2 = 0.999$ were used. For the training and validation of the GAN, we utilized an NVIDIA RTX 3090.

Top-level system configuration

As part of further investigations and optimizations of the Astrone exploration concept, we identified the challenge of increasing the resolution of a 128×128 pixels depth image to the resolution of the grayscale 2D camera image with a 1024×1024 pixels resolution. Our implemented approach consists of the trained GN without the DN for upscaling the depth image (green), pre-processing (orange), and post-processing (blue), as shown in Fig. 2:

1. *Pre-processing*: We rescale the low-resolution depth image and grayscale 2D camera image to relative values with a scale of [0,1] and create a sparse depth image from the low-resolution simulated LiDAR data (depth image) with the same resolution as the corresponding camera

image. Pixels with lacking information are initialized with 0. After, we concatenate the camera and depth image to get $1024 \times 1024 \times 2$ shaped data, which we divide into 16 non-overlapping $256 \times 256 \times 2$ image tiles.

2. *Depth image upscaling*: The GN increases the resolution of the sparse depth image by completing the missing values. The GN processes each $256 \times 256 \times 2$ tile separately.
3. *Post-processing*: The 16 GN's 256×256 output tiles per complete image are mosaiced to generate the expected high-resolution 1024×1024 depth image.

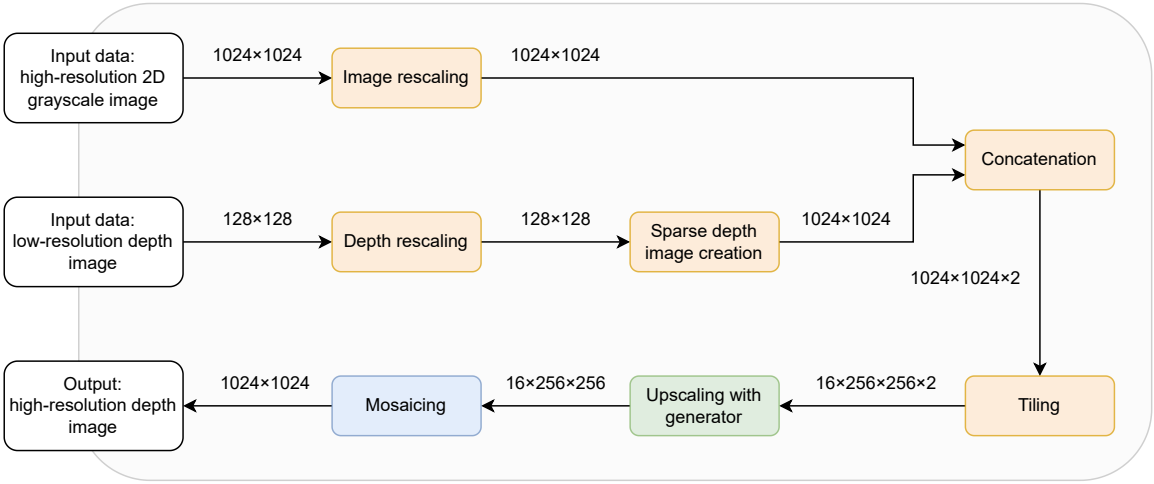


Figure 2. System processing including data dimension with pre- (orange) and post-processing (blue) subsystems, and the GAN's trained generator network for sparse depth image upscaling (green)

EXPERIMENTS AND RESULTS

Overview and metrics

The trained GN has been validated with a separate test dataset, that was not used to train either the GN or the DN. The test dataset includes challenges, such as irregular terrain, wide-angle optics, discontinuities at stone edges, and shadows caused by different lighting conditions. Subsequently, the GN's performance has been further investigated for different sun azimuths. The following metrics, which are described in Eq. 5a - 5d, evaluate the performance of our proposed approach. These metrics produce quantitative results by comparing the ground truth $\mathbf{y}$ with the system's predicted high-resolution depth image $\hat{\mathbf{y}}$. Smaller values represent better results for Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). We also compute the Peak-Signal-to-Noise Ratio (PSNR) using the maximum value $\|\mathbf{y}\|^2_{max}$ of the ground truth and the Mean Squared Error (MSE) between $\mathbf{y}$ and $\hat{\mathbf{y}}$. A higher PSNR represents a more accurate depth values reconstruction of the system's prediction. Additionally, the Structural Similarity Index Measure (SSIM) measures the similarity by considering their luminance, contrast, and, specifically here, important structural information. The SSIM's output values are between -1 and 1, where values closer to 1 represent more accurate depth value reconstructions of the system. The SSIM is calculated from the means $\mathbb{E}_{\hat{\mathbf{y}}_k}$, $\mathbb{E}_{\mathbf{y}_k}$ and their variances σ. The values of C_1 and C_2 have been selected as $C_1 = 0.01$ and $C_2 = 0.03$, respectively.

It is important to note that the average values of these metrics are calculated based on the number of test samples n. Further, a scaling factor of 100 was applied to MAE and RMSE for better visualization.

$$MAE = \frac{100}{n} \cdot \sum_{k=1}^{n} |\hat{\mathbf{y}}_k - \mathbf{y}_k| \tag{5a}$$

$$RMSE = \frac{100}{n} \cdot \sum_{k=1}^{n} \sqrt{\frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} (\hat{\mathbf{y}}_{k,i,j} - \mathbf{y}_{k,i,j})^2} \tag{5b}$$

$$PSNR = \frac{1}{n} \cdot \sum_{k=1}^{n} 10 \cdot \log \left(\frac{\|\mathbf{y}\|_{max}^2}{MSE} \right) \tag{5c}$$

$$SSIM = \frac{1}{n} \cdot \sum_{k=1}^{n} \frac{(2\mathbb{E}_{\hat{\mathbf{y}}_k}\mathbb{E}_{\mathbf{y}_k} + C_1)(2\sigma_{\hat{\mathbf{y}}_k\mathbf{y}_k} + C_2)}{(\mathbb{E}_{\hat{\mathbf{y}}_k}^2 + \mathbb{E}_{\mathbf{y}_k}^2 + C_1)(\sigma_{\hat{\mathbf{y}}_k}^2 + \sigma_{\mathbf{y}_k}^2 + C_2)} \tag{5d}$$

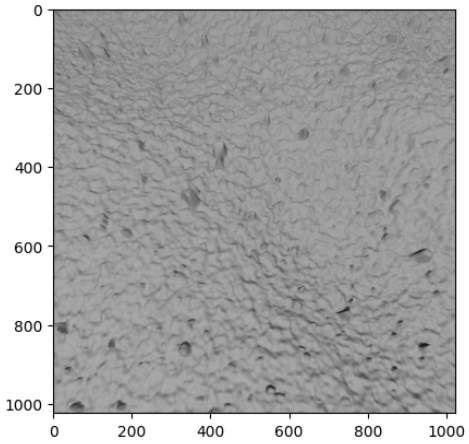
Validation and Testing

This section conducts experiments to validate our model's effectiveness in upscaling low-resolution depth images. We applied the workflow presented in the previous section to grayscale 2D images and the corresponding low-resolution depth images, from our synthetic test dataset, as the system's input data. Fig. 3 shows the input data for one sample: the grayscale 2D image, the low-resolution depth image, the upscaled high-resolution depth image, and a comparison to the ground truth. Visually, there are practically no differences between the high-resolution depth image (Fig. 3(c)) and ground truth (Fig. 3(d)). Therefore, Fig. 3(e) depicts the absolute error $|\mathbf{y}–\hat{\mathbf{y}}|$ for a better comparison. From this image, it is visible that the highest errors are around the edges of different-sized stones. In comparison, the error on the irregular surface is relatively small. The visualization also makes it possible to see a higher error at the verges of each tile due to our necessary preprocessing. Otherwise, despite using a large FoV, problems due to wide-angle effects are not visible in both the output and visualized error. In Fig. 3, each axis was labeled with the corresponding pixels to highlight the increase in resolution by a factor of 8×8.

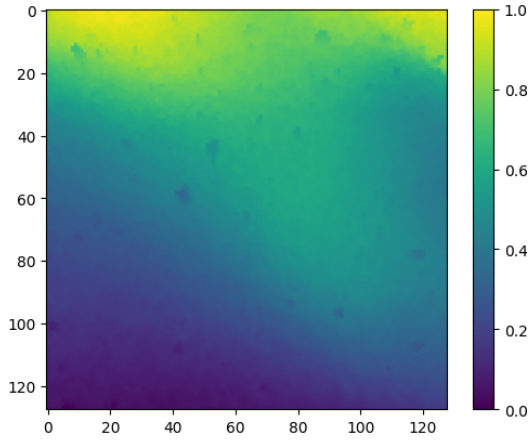
Further, we validated our model with $3,000$ individual test samples for a quantitative evaluation. As a reference, we used the same data to upscale our low-resolution depth image input with a bicubic interpolation as a widely known and more traditional approach. The sun azimuth angle of 0° was given as a nominal condition. Therefore no shadows are visible in the high-resolution grayscale 2D image. Tab. 1 includes the generated test results. For comparison, we applied the metrics mentioned before to the output data. Our approach achieves smaller error values and a higher PSNR. Even the SSIM shows an improvement.

Table 1. Quantitative Comparison on our test dataset for a sun azimuth of 0°

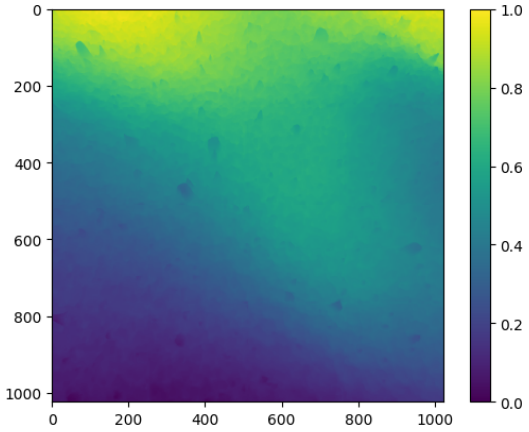
Method	MAE	RSME	PSNR (dB)	SSIM
Interpolated low-resolution data	0.626	0.876	41.110	0.964
Proposed	0.409	0.552	45.465	0.994



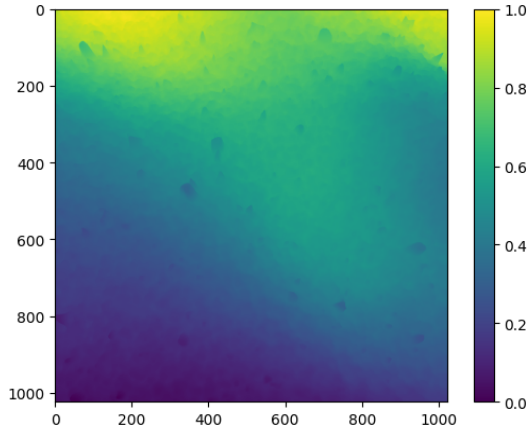
(a) High-resolution 2D grayscale image



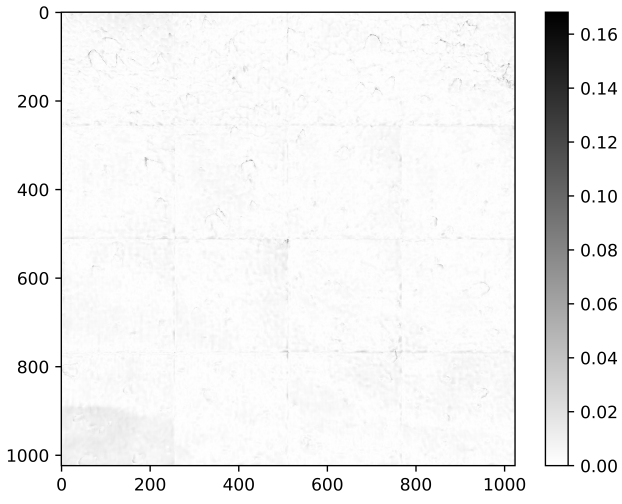
(b) Low-resolution depth image



(c) Prediction



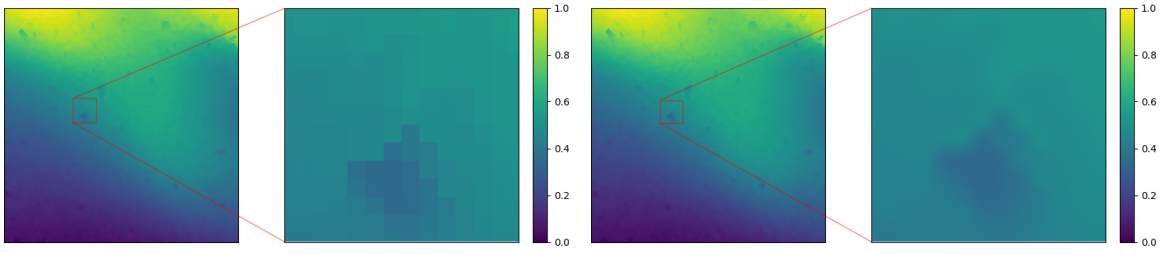
(d) Ground truth



(e) Error

Figure 3. Visualization of the input, output, ground truth, and the absolute error between the prediction and ground truth

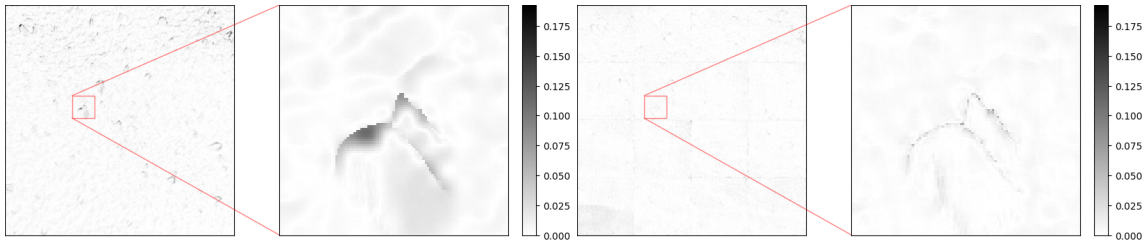
Fig. 4 presents a more detailed difference comparison of the original low-resolution depth image, the bicubic interpolated low-resolution data, the result of one sample of our method, and the ground truth. The interpolated low-resolution depth image (Fig. 4(b)) produces a blurry output and lacks in detail. Individual objects cannot be detected. Our GAN-based approach generates a much sharper and more detailed result (Fig. 4(c)). The two stones in the foreground can be distinguished. The higher level of detail is visible, especially around the edges of stones or boundaries between stones and the background. The overall output is close to the ground truth. To visualize the error of the presented approach, Fig. 5 shows the absolute error.



(a) Low-resolution (b) Interpolated low-resolution data

(c) Prediction (d) Ground truth

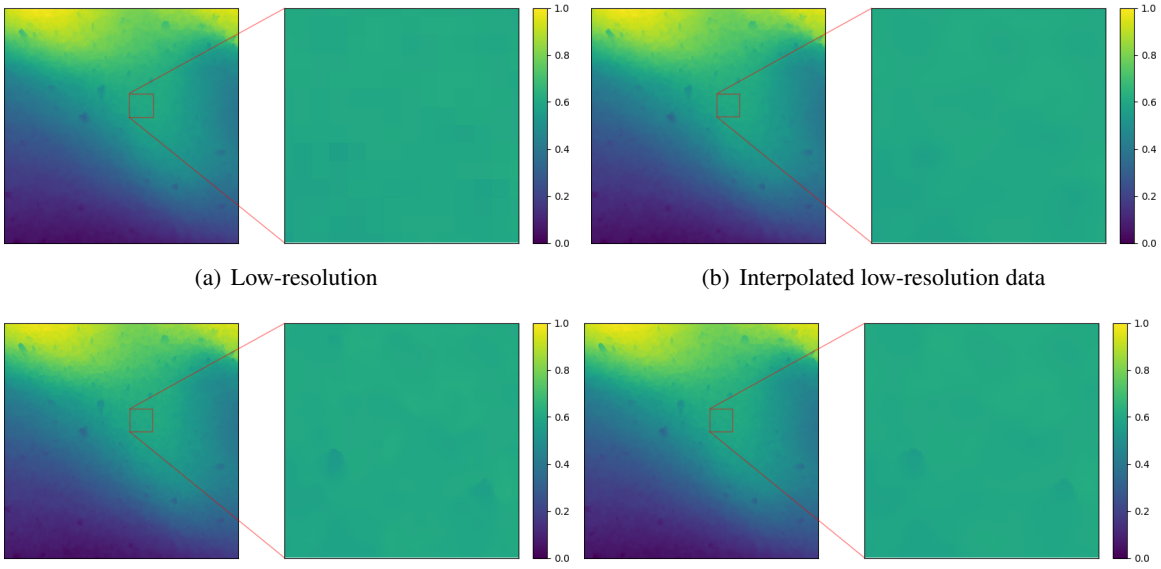
Figure 4. Visualization of the reconstruction of bigger objects



(a) Interpolated low-resolution data (b) Prediction

Figure 5. Visualized absolute error between ground truth, the interpolated low-resolution data and the prediction of big stones

Fig. 6 shows the difference in regard to small object detection. Small objects are already hardly detectable or almost disappear in the original low-resolution depth image. This is also true after upscaling this data with interpolation. Our method is still able to reconstruct these smaller objects, which are similar to the expected ground truth. This was achieved even when the low-resolution input depth image provided no longer sufficient information for a rough stone shape.



(a) Low-resolution (b) Interpolated low-resolution data

(c) Prediction (d) Ground truth

Figure 6. Visualisation of the reconstruction of small objects

Intuitively, the output of our method is close to the ground truth depth images, especially in terms of perception and restoration for edges and contours. Edges are generally sharper than in low-resolution data. It is possible to detect individual objects and ensure the preservation of smaller ones. The qualitative results indicate that the GN uses high-frequency components from the grayscale 2D camera images to restore these details.

Our methodology employs simulated high-resolution monocular grayscale 2D camera images. Therefore, the second experiment was aimed at assessing our trained GN's robustness to changing illumination conditions. According to the additional generated data with an altering sun azimuth angle in $10°$ increments within a range of $\pm 80°$ in the sensor frame, the corresponding conditions were tested. The test compared the resulting predictions against the ground truth for $1,000$ images for each angle. All metrics are computed for the 1024×1024 output rather than just one tile. It should be noted that the model was not trained for different illumination conditions. According to the data presented in Tab. 2, it has been observed that the sun's azimuth influences the model's predictions as expected. However, this influence is relatively small until more extreme shadows occur. Furthermore, whether the sun azimuth is negative or positive has a negligible impact on the outcomes. The results of this test show that the GAN's GN uses the high-resolution grayscale 2D camera image for information extraction here too. Fig. 7 provides a visual representation of the grayscale 2D camera images, predictions, and absolute error for sun azimuth angles of $0°$, $30°$, and $80°$, resulting in different levels of shadows. A close qualitative inspection of the visualized absolute error shows that the overall error of the output increases as the size and amount of shadows increase. Despite the presence of more shadows, effects due to the wide-angle lens used could not be observed. These results show that our trained GN has some degree of robustness against varying illumination conditions across the frame, even without specific training, for more shadow robustness.

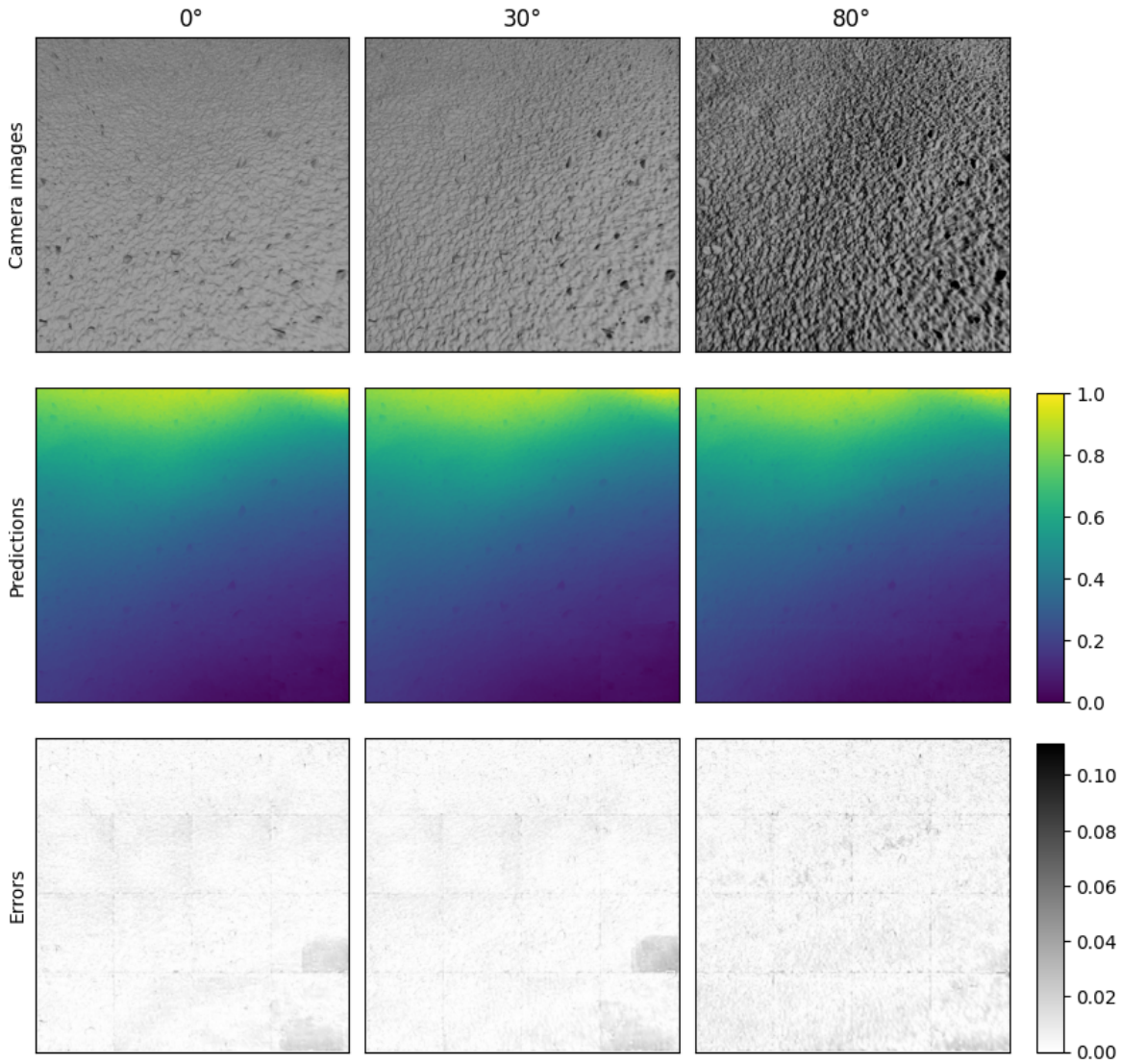


Figure 7. Comparison of the predictions (second row) for grayscale images with different sun azimuths (first row) and the visualized absolute error between ground truth and the prediction (third row)

Computational Time

The computational time of the system includes the pre- and post-processing stages, along with the execution process of the GN and the inference time of the GN itself. For every pair of low-resolution depth and high-resolution grayscale 2D images, the neural network processes 16 corresponding tiles. We utilize identical hardware for testing and training: a Ryzen 9 5950x CPU and an NVIDIA RTX 3090 GPU. Using the GPU for the computations of the GN, it took $8.47\,\mathrm{s}$ on average out of 100 runs. However, the processing still needs to be optimized explicitly for particular target hardware, which provides good optimization opportunities for future runtime optimization. For the specific mission scenario involving the Astrone spacecraft, the execution time is not critical due to the spacecraft's extremely slow dynamics on the comet 67P/Churyumov–Gerasimenko's surface.

Table 2. Validation on different sun azimuth

sun azimuth	MAE	RSME	PSNR (dB)	SSIM
$-80°$	0.559	0.769	42.57	0.986
$-70°$	0.515	0.714	43.22	0.987
$-60°$	0.484	0.675	43.74	0.988
$-50°$	0.460	0.645	44.18	0.989
$-40°$	0.441	0.618	44.58	0.991
$-30°$	0.424	0.594	44.98	0.992
$-20°$	0.414	0.577	45.27	0.993
$-10°$	0.414	0.575	45.36	0.993
$10°$	0.414	0.575	45.35	0.993
$20°$	0.414	0.578	45.27	0.993
$30°$	0.424	0.594	44.98	0.992
$40°$	0.440	0.618	44.58	0.991
$50°$	0.460	0.644	44.20	0.989
$60°$	0.484	0.675	43.74	0.988
$70°$	0.515	0.714	43.23	0.987
$80°$	0.558	0.769	42.58	0.986

CONCLUSION

In this paper, we propose an AI-based approach to increase the resolution of a low-resolution depth image by data fusion with a single monocular 2D grayscale image close to the irregular and unstructured surface of SSSBs. We introduced a GAN-based method to ensure the reconstruction of high-frequency components and small details. A neural network was trained and validated with a realistic synthetic dataset containing $10,000$ samples. Additionally, data with different sun positions were created to simulate changes in local illumination during a real mission. The validation successfully demonstrated the ability to accurately increase the resolution of coarse flash-LiDAR data by a factor of 8×8. The sharper and more detailed output made it possible to detect individual and small objects that were not distinguishable or visible in the low-resolution flash-LiDAR data. Furthermore, robustness to changing illumination conditions has been shown despite the network not being trained specifically for these cases. Also, no effects of wide-angle imaging ($65°\times65°$) were observed during testing.

The presented solution shows the potential for AI-based approaches to computer vision tasks in the context of near-surface navigation applications on irregular SSSBs.

ACKNOWLEDGMENT

The project “Astrone KI - AI support for high surface mobility of planetary research platforms (agile, autonomous, robust).” has received funding from the German Federal Ministry for Economic Affairs and Energy (BMWi) under funding number “50 RA 2130C” supervised by the German Space Agency (DLR).

REFERENCES

[1] G. Ono, F. Terui, N. Ogawa, S. Kikuchi, Y. Mimasu, K. Yoshikawa, H. Ikeda, Y. Takei, S. Yasuda, K. Matsushima, *et al.*, "GNC strategies and flight results of Hayabusa2 first touchdown operation," *Acta Astronautica*, Vol. 174, 2020, pp. 131–147.

[2] D. Lauretta, S. Balram-Knutson, E. Beshore, W. Boynton, C. Drouet d'Aubigny, D. DellaGiustina, H. Enos, D. Golish, C. Hergenrother, E. Howell, *et al.*, "OSIRIS-REx: sample return from asteroid (101955) Bennu," *Space Science Reviews*, Vol. 212, 2017, pp. 925–984.

[3] P. Michel, M. Kueppers, H. Sierks, I. Carnelli, A. F. Cheng, K. Mellab, M. Granvik, A. Kestilä, T. Kohout, K. Muinonen, *et al.*, "European component of the AIDA mission to a binary asteroid: Characterization and interpretation of the impact of the DART mission," *Advances in Space Research*, Vol. 62, No. 8, 2018, pp. 2261–2272.

[4] M. Martin, F. Belien, B. Liu, J. Olucak, F. Schimpf, T. Brüggemann, A. Falke, V. Sazdovski, V. Chernykh, K. Janschek, *et al.*, "Astrone–GNC for Enhanced Surface Mobility on Small Solar System Bodies," *11 th International ESA Conference on Guidance, Navigation & Control Systems*, 2021.

[5] B. Liu, V. Sazdovski, V. Chernykh, and K. Janschek, "Flash LiDAR Aided-Inertial Navigation on Surfaces of Small Solar System Bodies Using Error State Kalman Filtering," *AAS/AIAA Astrodynamics Specialist Conference*, 2021. Paper AAS 21-752.

[6] R. W. Gaskell, O. S. Barnouin-Jha, D. J. Scheeres, A. S. Konopliv, T. Mukai, S. Abe, J. Saito, M. Ishiguro, T. Kubota, T. Hashimoto, J. Kawaguchi, M. Yoshikawa, K. Shirakawa, T. Kominato, N. Hirata, and H. Demura, "Characterizing and Navigating Small Bodies with Imaging Data," Vol. 43, No. 6, pp. 1049–1061, 10.1111/j.1945-5100.2008.tb00692.x.

[7] M. M. A. Asad, L. C. Philpott, C. L. Johnson, O. S. Barnouin, E. Palmer, J. R. Weirich, M. G. Daly, M. E. Perry, R. Gaskell, E. B. Bierhaus, J. A. Seabrook, R. Espiritu, H. Nair, C. Ernst, R. T. Daly, M. C. Nolan, H. L. Enos, and D. S. Lauretta, "Validation of Stereophotoclinometric Shape Models of Asteroid (101955) Bennu during the OSIRIS-REx Mission," Vol. 2, No. 2, 2021, p. 82, 10.3847/PSJ/abe4dc.

[8] Y. Ming, X. Meng, C. Fan, and H. Yu, "Deep learning for monocular depth estimation: A review," *Neurocomputing*, Vol. 438, May 2021, pp. 14–33, 10.1016/j.neucom.2020.12.089.

[9] J. Hu, C. Bao, M. Ozay, C. Fan, Q. Gao, H. Liu, and T. L. Lam, "Deep Depth Completion: A Survey," May 2022, 10.48550/arXiv.2205.05335.

[10] Z. Wang, J. Chen, and S. C. H. Hoi, "Deep Learning for Image Super-Resolution: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 43, Oct. 2021, pp. 3365–3387. Number: 10 Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, 10.1109/TPAMI.2020.2982166.

[11] J. Zhao, Y. Wang, Y. Cao, M. Guo, X. Huang, R. Zhang, X. Dou, X. Niu, Y. Cui, and J. Wang, "The Fusion Strategy of 2D and 3D Information Based on Deep Learning: A Review," *Remote Sensing*, Vol. 13, Oct. 2021, p. 4029. Number: 20, 10.3390/rs13204029.

[12] Z. Xie, X. Yu, X. Gao, K. Li, and S. Shen, "Recent Advances in Conventional and Deep Learning-Based Depth Completion: A Survey," *IEEE Transactions on Neural Networks and Learning Systems*, 2022, pp. 1–21, 10.1109/TNNLS.2022.3201534.

[13] Z. Chen, B. Wu, and W. C. Liu, "Mars3DNet: CNN-Based High-Resolution 3D Reconstruction of the Martian Surface from Single Images," Vol. 13, No. 5, p. 839, 10.3390/rs13050839.

[14] H. Chen, X. Hu, and J. Oberst, "PIXEL-RESOLUTION DTM GENERATION FOR THE LUNAR SURFACE BASED ON A COMBINED DEEP LEARNING AND SHAPE-FROM-SHADING (SFS) APPROACH," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. V-3-2022, May 2022, pp. 511–516, 10.5194/isprs-annals-V-3-2022-511-2022.

[15] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," , 10.48550/arXiv.1406.2661

[16] Y. Tao, J.-P. Muller, S. Xiong, and S. J. Conway, "MADNet 2.0: Pixel-Scale Topography Retrieval from Single-View Orbital Imagery of Mars Using Deep Learning," Vol. 13, No. 21, p. 4220, 10.3390/rs13214220.

[17] Y. Liu, Y. Wang, K. Di, M. Peng, W. Wan, and Z. Liu, "A Generative Adversarial Network for Pixel-Scale Lunar DEM Generation from High-Resolution Monocular Imagery and Low-Resolution DEM," Vol. 14, No. 21, p. 5420, 10.3390/rs14215420.

[18] R. La Grassa, I. Gallo, C. Re, G. Cremonese, N. Landro, C. Pernechele, E. Simioni, and M. Gatti, "An Adversarial Generative Network Designed for High-Resolution Monocular Depth Estimation from 2D HiRISE Images of Mars," Vol. 14, No. 18, p. 4619, 10.3390/rs14184619.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, Springer, 2015, pp. 234–241.

[20] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation," Apr. 2020.

[21] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," Nov. 2018.

[22] J. Eggert, S. Weikert, and I. Kossev, "Imaging Sensor Emulation and Dynamics Simulation for PIL/HIL,"

[23] Blender Online Community, *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018.