



Article

A Generative Adversarial Network for Pixel-Scale Lunar DEM Generation from High-Resolution Monocular Imagery and Low-Resolution DEM

Yang Liu ^{1,2}, Yexin Wang ^{1,*}, Kaichang Di ^{1,3} , Man Peng ¹, Wenhui Wan ¹ and Zhaoqin Liu ¹

¹ State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

² University of Chinese Academy of Sciences, Beijing 100049, China

³ CAS Center for Excellence in Comparative Planetology, Hefei 230026, China

* Correspondence: wangyx@aircas.ac.cn; Tel.: +86-10-64807987

Abstract: Digital elevation models (DEMs) provide fundamental data for scientific and engineering applications in lunar exploration missions. Lunar DEMs have been mainly generated by laser altimetry and stereophotogrammetry. Complementarity to stereo photogrammetry, reflection-based surface reconstruction methods, such as shape from shading (SFS), have been studied and applied in lunar DEM reconstruction from a single image. However, this method often suffers from solution ambiguity and instability. In this paper, we propose a generative adversarial network (GAN)-based method that is able to generate high-resolution pixel-scale DEMs from a single image aided by a low-resolution DEM. We have evaluated the accuracy of the reconstructed high-resolution DEMs from 25 LROC NAC images of four regions using LROC NAC DEMs (2 m/pixel) as ground truth. The experimental results demonstrate good accuracy and adaptability to changes in illumination conditions. The root mean square error (RMSE) can reach about 2 m in areas where the degree of elevation variation is less than 100 m, and the RMSE value ranges from around 3 m to 10 m without considering the degree of the elevation variation in large-area reconstruction. As high-resolution monocular images and low-resolution DEMs are available for the entire lunar surface, the proposed GAN-based method has great potential in high-resolution lunar DEM reconstruction for lunar mapping applications.

Keywords: lunar exploration; digital elevation model; high-resolution imagery; generative adversarial network



Citation: Liu, Y.; Wang, Y.; Di, K.; Peng, M.; Wan, W.; Liu, Z. A Generative Adversarial Network for Pixel-Scale Lunar DEM Generation from High-Resolution Monocular Imagery and Low-Resolution DEM. *Remote Sens.* **2022**, *14*, 5420. <https://doi.org/10.3390/rs14215420>

Academic Editor: Roberto Orosei

Received: 14 September 2022

Accepted: 25 October 2022

Published: 28 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Digital elevation models (DEM) of the Moon provide the fundamental data in lunar topographic, morphological, and geological studies. They are also essential in engineering applications in lunar exploration missions, e.g., in landing site evaluation, safe landing, rover navigation, etc. [1–6]. Lunar global DEMs have been produced using laser altimetry or optical images. Representative examples include the lunar orbiter laser altimeter (LOLA) LDEM1024 with a resolution of 1024 pixels per degree (ppd) (~30 m at the equator) [7], 1024 ppd DEM derived from 10 m-resolution SELENE terrain camera (TC) stereo images [8], and 20 m grid spacing DEM derived from 7 m-resolution Chang'e-2 stereo images [9]. SLDEM2015, which was produced by coregistration and a combination of SELENE TC DEM and LOLA data, is also widely used. It covers latitudes within $\pm 60^\circ$ at a horizontal resolution of 512 ppd (~60 m at the equator) [10].

Lunar local DEMs with higher resolution (e.g., meter-level resolution) have also been produced, mostly by stereo photogrammetry, which requires target area covering by stereo images. As a rule of thumb, the resolution (grid spacing) of a DEM generated by stereophotogrammetry is about three times that of the stereo imagery [11]. Currently,

the highest-resolution lunar orbital images (up to 0.5 m/pixel) are acquired by lunar reconnaissance orbiter camera (LROC) narrow-angle camera (NAC), and the local DEMs generated from LROC NAC stereo images are usually of 2–5 m resolution [12], which are the highest-resolution DEMs derived from orbital images. The LROC was not designed with built-in stereo observation capability, and NAC stereo images can only be acquired from neighboring orbits using off-nadir slew. As a result, NAC stereo observations are very limited, though NAC monocular images have covered almost the entire lunar surface. Therefore, it is highly desirable and will be widely applicable to reconstruct pixel-scale DEMs from monocular high-resolution orbital images, particularly LROC NAC images.

In fact, complementary to stereo photogrammetry, shape from shading (SFS) (also called photoclinometry) techniques have been studied and applied in the planetary mapping field [11,13–15], which reconstruct a DEM from a single image or multiple nonstereo images. SFS techniques have also been widely researched in the field of computer vision and artificial intelligence [16,17]. In general, the multiple-image SFS method is limited by a lack of multiple images with different illumination conditions covering the same area. According to [18], single-image SFS is underconstrained and faces serious issues. First, the assumption of a uniform or known albedo is needed, but determining the albedo from a single image is underconstrained. Second, as shading actually reflects the depth of a point relative to its neighbors, it is difficult to get clues to restore shapes from coarse-scale depth variations that produce small changes throughout the entire image. A number of approaches have been developed to overcome the underconstrained issues, including learning albedos from training data [18], separation of albedos, and shape optimization with the aid of coarse DEMs, etc. [15,19].

With the development of deep learning, a neural network can learn the “hidden” mapping relationship between two types of data and realize the “end-to-end” transformation of data under the condition of sufficient training samples. Deep neural network has been applied to DEM reconstruction using a single image and has shown some advantages in establishing reliable mapping and eliminating noise and image uncertainty. In order to improve the performance of DEM reconstruction, strategies of deepening the depth of the network and increasing the amount of training samples can be used [20]. There is no atmospheric influence on the lunar surface, and the change of image intensities is mainly caused by the topography, supposing that the albedo does not change severely in a local area. Low-resolution DEMs are available for the entire lunar surface and can be used as a constraint to enable the network to obtain low-frequency terrain information, while high-frequency terrain information can be obtained from high-resolution image. This makes it possible to use a network to generate a high-resolution DEM with higher accuracy from high-resolution monocular imagery and existing low-resolution DEM.

In this paper, a high-resolution lunar DEM reconstruction method is proposed based on generative adversarial network (GAN), which generates pixel-scale lunar DEM from monocular LROC NAC imagery aided by existing low-resolution DEM, e.g., SLDEM. For convenience, we call the lunar DEM reconstruction network LDEMGAN. It adopts the U-Net structure as the generator network, uses the convolution layer instead of the pooling layer for feature extraction, takes the discrimination result of the discriminator network instead of the image similarity as the loss, adds the gradient loss item, and uses a variety of training data to train the network. In this study, four regions are selected for reconstruction testing using 25 LROC NAC image (2 m/pixel) and SLDEM (~60 m/pixel). The reconstruction results are verified with DEMs (2 m/pixel) produced by LROC NAC images, and the effects of different lunar surface conditions and data quality on the reconstruction results are analyzed. The effects of adding low-resolution DEM and influence of illumination conditions of the images are also discussed.

Section 2 of this paper reviews the relevant research. Section 3 presents the proposed method in detail. Section 4 gives the experimental results and discussions. Conclusions are given in Section 5.

2. Related Work

With rapid development of deep learning technology, monocular depth estimation approaches based on deep neural networks have been studied and applied in various fields of computer vision, such as robotics, path planning, autonomous navigation, etc. Due to difficulties of absolute depth recovery, as well as the demand of comparable accuracy to that of traditional triangulation methods, monocular depth estimation is still an active and challenging topic. Currently, the commonly used deep neural network models for monocular depth recovery are the convolutional neural network (CNN), recurrent neural network (RNN), and generative adversarial network (GAN) [21]. As RNN is usually employed in the field of video sequence processing, this section only focuses on reviewing and discussing monocular depth restoration that uses CNN and GAN models.

CNN models have been well developed and applied in monocular depth estimation. Eigen et al. [22] adopted a two-scale convolutional network to first estimate the depth of the scene on a global scale and then add details on a local scale. The work also raised and explained the basic problem of solving monocular depth estimation in terms of supervised learning [22]. On this basis, Eigen et al. subsequently proposed an improved framework constructed by three-scale network structure for monocular depth estimation [23]. Thereafter, Laina et al. [24] proposed a fully convolutional network structure to capture monocular features at multiple scales with fewer parameters, and solved the problem of high training costs for fully connected layers. An up-convolution decoding module was designed to improve the processing speed of the network [24]. Ma et al. proposed a CNN-residual network, and the comparison to the network with the hybrid convolutional fully connected network and to the fully convolutional network demonstrated the effectiveness and optimality of the proposed network [25]. Fu et al. added a spacing-increasing discretization strategy that splits the depth according to different intervals and recast the network learning as an ordinal regression problem to reduce the training loss weight in regions with large depth values [26]. Facil et al. used the U-Net to encode images and then add sensor parameters for decoding, which enhances the generalization ability of the network [27]. In order to shorten the computing time of the network, Wolk et al. proposed a lightweight U-Net to improve the operation speed, which can realize real-time depth estimation [28].

The GAN model was proposed by Goodfellow in 2014 and has been proved applicable for monocular depth estimation tasks [29]. Jung et al. used GAN for the first time to implement monocular depth estimation by supervised learning [30]. The generative network consists of a global-scale U-Net [31] and a fully convolutional fine-scale network. The discriminative network is designed according to Li's work [32]. Instead of using a multiscale generator, Lore et al. used two GANs for depth estimation at both global and local scales [33].

In the field of planetary remote sensing, since the orbital images of the Moon's surface, especially high-resolution stereo images, are very scarce compared to those of the Earth, DEM generation by monocular high-resolution images is needed more than traditional photogrammetry methods. Therefore, in recent years, several studies have used depth estimation networks to reconstruct relative heights from planetary orbital images and to restore relative heights to absolute heights using control points or existing coarse DEM. Several works have discussed combining monocular relative height estimation and scale information recovery for planetary surface reconstruction. Chen et al. proposed a CNN-based digital terrain model (DTM) reconstruction method for Mars using Mars Reconnaissance Orbiter Context Camera (CTX) images [20]. Tao et al. proposed a GAN-based network called MADNet for fast reconstruction of DTM from a single image of Mars [34]. Subsequently, MADNet 2.0 was developed, which used the optimized single-scale inference and multiscale reconstruction to improve the resolution of reconstructed DTM [35]. These methods all used a two-step strategy to reconstruct the terrain, i.e., firstly, an image is input into the deep network to restore relative depth, and then absolute height is recovered by stretching the relative depth using a scale value calculated by existing low-resolution data.

In these previous works, low-resolution data are only used for scale recovery such that the reconstruction result is affected greatly by scale value. To make full use of the coarse DEM, Hao et al. proposed using the low-resolution DEM and the image as the input of the CNN, and the CNN prediction results and the image obtain a high-resolution DEM by SFS [36]. However, the reconstruction results generated by SFS are affected by lighting conditions. Intending to make full use of low-resolution DEM in terrain reconstruction, improve adaptability to light conditions and combine the advantages of GAN, this paper proposes a GAN-based reconstruction network for lunar DEM generation—LDEMGAN. The main features of this work are:

- (1) The proposed approach takes both monocular image and a coarse DEM as input of the GAN.
- (2) The coarse DEM is fully used to provide control information in the proposed LDEMGAN so as to improve the stability and accuracy of high-resolution DEM reconstruction.
- (3) In order to generate DEM at the same high resolution as the input image, the reconstruction process can be executed iteratively until the needed DEM resolution is obtained.

3. Method

3.1. Overview of the Proposed Approach

In this study, LROC NAC imagery is used to reconstruct a pixel-scale DEM through LDEMGAN with the help of control information provided by SLDEM. Figure 1 shows the overall process of the proposed approach, which consists of training and reconstruction phases. In the beginning, the registered LROC NAC imagery and SLDEM data are preprocessed. (1) The centers of the craters with diameters of about 200 m are selected as control points, and the LROC NAC images and SLDEM are registered by affine transformation. The registration accuracy needed to reach the subpixel level of SLDEM (the registration error needs to be less than the grid spacing of SLDEM). (2) The NAC image performs downsampling if needed and the SLDEM is upsampled to the same pixel size. (3) They are both clipped to the size of 256×256 due to the requirement of the network with an overlap of 100 pixels between adjacent image blocks. (4) The clipped data are normalized to $[0, 1]$. (5) The clipped image and SLDEM data are set as channel 1 and channel 2, respectively, to synthesize the input data of LDEMGAN. Note that in the training phase, the corresponding NAC DEM is also preprocessed in the same way so as to take the role as ground truth.

In the training phase, the preprocessed high-resolution image and low-resolution DEM of the same area with the same grid are taken as input of LDEMGAN, which outputs depth reconstruction results. The ground truth is preprocessed DEM with identical resolution of the input imagery, and supervised training is carried out to obtain the reconstruction network.

In the reconstruction phase, the preprocessed high-resolution image and low-resolution DEM covering the same area are input into LDEMGAN to generate DEM that has identical resolution to the input image. However, in practice, the resolutions of the input image and the DEM that provide control information differ too much, such as NAC and SLDEM in our case. Therefore, a high-resolution DEM generation process can be carried out iteratively, which is shown in Figure 1. In our case, first the low-resolution SLDEM is upsampled to 10 m grid spacing and an input high-resolution NAC image is downsampled to the same pixel size of the upsampled SLDEM. Next, they are simultaneously input into the network to generate DEM with identical resolution to the downsampled input image which is 10 m/pixel. As the resolution of the generated DEM is lower than the NAC image, the previous steps should be implemented again. This time, the generated DEM from the previous step is upsampled to the same pixel size of the NAC image. Then, they are both input into the reconstruction network to obtain a DEM with the same resolution as the NAC image. Afterwards, scale restoration will be carried out to convert the generated DEM from relative height into absolute height.

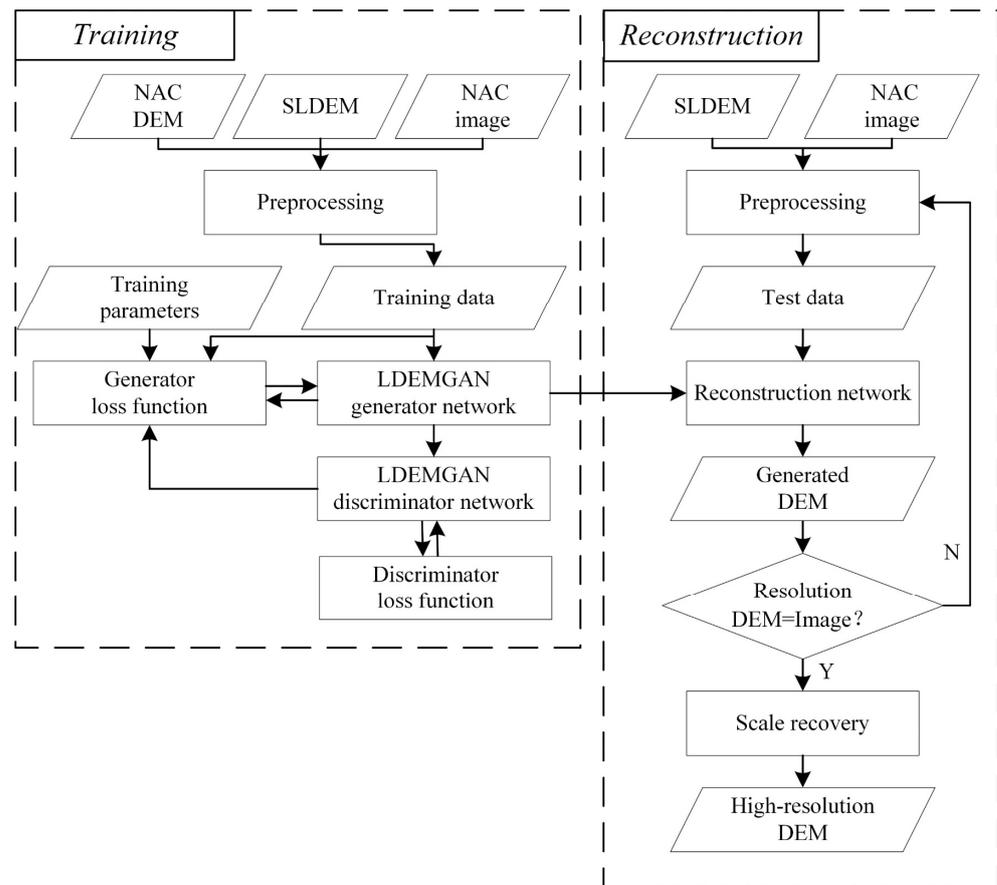


Figure 1. Overall process of the proposed method.

It is worth mentioning that there are often geometric inconsistencies between data captured from different sensors; therefore, geometric registration of the NAC image and SLDEM is required beforehand in this study. In fact, the accuracy of geometric registration can directly affect the final reconstruction result.

3.2. Network Architecture

Figure 2 shows the LDEMGAN network architecture used for high-resolution DEM generation. According to the principle of the GAN network, it can be divided into a generator network G and a discriminator network D. The generator network adopts a U-Net architecture, adding skip connections between the encoding part and the decoding part. The CNN structured encoder consists of 7 BCDL modules, which encode a 256×256 image into a feature vector with a size of $[2, 2, 1024]$. The BCDL module consists of a batch normalization layer, a convolution layer (1×1 kernel, and with increasing number of feature maps of 64, 128, 256, 512, 1024, 1024, 1024, stride 1), a downsampling convolution layer (4×4 kernel, stride 2, padding 1), and a LeakyReLU activation function. LeakyReLU is able to solve the zero gradient problem for negative values, as shown in Equation (1):

$$\text{LeakyReLU}(x) = \begin{cases} x & x > 0 \\ 0.2x & x \leq 0 \end{cases} \quad (1)$$

where x is the input tensor. The decoder employs a BUR model and adds skip connections, enabling direct transfer of low-frequency information. The BUR module consists of a batch normalization layer, an upsampling convolution layer (4×4 kernel, and with decreasing

number of feature maps of 1024, 1024, 512, 256, 128, 64, 1, stride 2, padding 1), and an activation function. The activation function used for the BUR-1–BUR-6 is ReLU:

$$\text{ReLU}(x) = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (2)$$

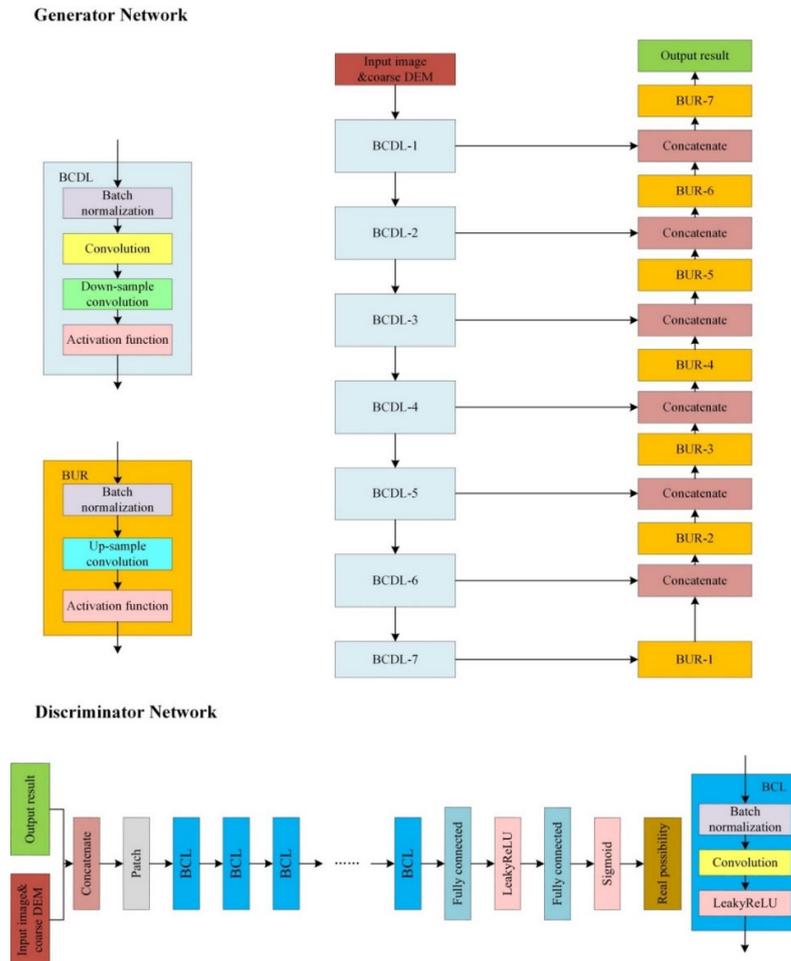


Figure 2. Schematic diagram of the LDEMGAN structure.

The TanhL activation function is used in the BUR-7, so that the range of the output result can be constrained within [0, 1], as shown below:

$$\text{TanhL} = \frac{1}{2} \left(\frac{e^x - e^{-x}}{e^x + e^{-x}} + 1 \right) \quad (3)$$

The optimization objective function of training the generator network parameters is as follows:

$$\hat{\theta}_G = \operatorname{argmin}_{\theta_G} \frac{1}{N} \sum_{n=1}^N L_G(H_{pred}^n, H_{truth}^n) \quad (4)$$

Among them, N is the total number of images, $L_G(\cdot)$ is the loss function of the generator network G (see Section 3.3), θ_G is the parameter of the generator network. H_{truth}^n is the n th ground truth DEM data of the corresponding input image, and H_{pred}^n is the n th reconstruction results derived by the generator network. H_{pred}^n is calculated as:

$$H_{pred}^n = G_{\theta_G}(I_n, H_{coarse}^n) \quad (5)$$

where G_{θ_G} is the generator network with network layer parameters θ_G , I_n is the n th input high-resolution image, and H_{coarse}^n is the corresponding low-resolution DEM that provides control information in the same area.

According to the structure of the GAN network, the output of the discriminator network D is a probability value that denotes whether its input sample is real or fake. The discriminator network used in this paper adopts the principle of PatchGAN to capture the high-frequency information. PatchGAN divides the output of the last layer into M blocks, and discriminates each block individually:

$$P_H = \frac{1}{M} \sum_{m=1}^M D((I^m, H_{coarse}^m), H^m) \quad (6)$$

where H^m is the H_{pred}^n calculated by the generator network or the ground truth surface data H_{truth}^n . Therefore, the optimization objective for the adversarial network parameters is:

$$\hat{\theta}_D = \operatorname{argmin}_{\theta_D} \frac{1}{N} \sum_{n=1}^N L_D(I^n, H_{pred}^n, H_{truth}^n) \quad (7)$$

where $L_D(\cdot)$ is the loss function of the discriminator network D (see Section 3.3).

3.3. Loss Function

The loss function of the LDEMGAN consists of generator network loss and discriminator network loss. The losses of the generator network include the loss representing the difference between the prediction and the ground truth, the loss of local structure, and the loss of the discriminator. Usually, the L_2 loss function is used for the generator network, which reflects the cumulative squared difference between the value predicted by the model and the ground truth value. However, it can bring over smooth predictions. Hence, the Huber loss is used, as shown in Equation (8):

$$L_{Huber}(H_{pred}, H_{truth}) = \begin{cases} (H_{pred} - H_{truth})^2 / 2 & |H_{pred} - H_{truth}| < \delta \\ \delta |H_{pred} - H_{truth}| - \frac{1}{2} \delta^2 & otherwise \end{cases} \quad (8)$$

where δ is one fifth of the maximum of the difference of prediction and the ground truth. When the difference is less than δ , the Huber loss is consistent with L_2 . Otherwise, the Huber loss equals L_1 .

In order to maintain the relationship between each pixel and its adjacent pixels, a local structure loss is added, as shown below:

$$L_{grad} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [(\nabla_x(H_{pred}, H_{truth}))^2 + \nabla_y(H_{pred}, H_{truth})^2] \quad (9)$$

where M is the number of image rows and N is the number of image columns. By considering ∇_x and ∇_y , the x directional and y directional gradient differences between the predicted value and the ground truth calculated in local windows, respectively, the consistencies of the predicted value and the ground truth can be preserved on a local scale.

The loss of discriminator refers to the discrimination of the generated prediction data by the discriminator network in LDEMGAN, indicating the similarity between the generated prediction data and the ground truth, as shown below:

$$L_{GAN} = \log(1 - D((I, H_{coarse}), G(I, H_{coarse}))) \quad (10)$$

where I is the input image and H_{coarse} is the input DEM (low-resolution DEM).

Therefore, the overall loss of the generator network can be represented by:

$$L_G = \alpha L_{Huber} + \beta L_{grad} + \gamma L_{GAN} \quad (11)$$

where α , β , γ are the coefficients to balance the losses.

The loss function of the discriminator network is shown as:

$$L_D = \frac{1}{2} [\log(D((I, H_{coarse}), G(I, H_{coarse}))) + \log(1 - D((I, H_{coarse}), H_{truth}))] \quad (12)$$

3.4. Scale Recovery

As normalization of the input data is needed before entering the generator network for better convergence, the values of the network output are within [0, 1]. Thus, post-processing is required to recover the reconstruction results to the correct scale as elevation information. This study uses the normalized low-resolution DEM as a constraint, as shown in Equation (10). In addition, in the scale recovery process, statistics of this low-resolution DEM are used to restore the absolute elevation. First, mean and standard deviation of the low-resolution DEM are calculated. Then the reconstruction results are stretched to obtain the absolute heights, as shown in Equation (13). By adding geographic information to the restored absolute heights, scale recovery is finally completed.

$$DEM_{abs} = m + sd \cdot DEM_{pred} \quad (13)$$

In the above equation, DEM_{abs} is the absolute elevation value, m is the mean value of the low-resolution DEM, and sd is the standard deviation of the low-resolution DEM.

3.5. Network Training

As mentioned in Section 3.1, practically the DEM restoration is carried out iteratively. In our case, two reconstruction networks are employed and the structures of them are identical. In fact, the restoration processes are similar in each iteration, so theoretically one set of trained network parameters should work for both networks. However, in order to obtain high-precision results, there is no doubt that the two networks should be trained separately with training data the same as the actual situation.

Meanwhile, in order to enhance the universality of the network, the number of training samples should be sufficient so that the network can avoid fuzzy solution caused by the changes of illumination conditions or various landform characteristics. For adapting illumination changes, the training data can be expanded from two aspects. First, the existing training data can be rotated and flipped to simulate different sun azimuths. Second, images captured at different sun altitude angles of the same area that pair with the same corresponding coarse DEM should be added to enhance the network's adaptation to shadow properties. For adapting to changes of landform characteristics in various topographic regions (e.g., lunar mare or highlands), the network performance can always be improved if the training data are adequate. Generally, if the feature characteristics of the area to be processed are hugely different from the areas that used for training the networks, the recommended practice is to employ transfer training by using a small amount of data from the destination area.

4. Experimental Results

4.1. Reconstruction Network Training

As described in the above sections, two networks with identical structure and trained by different data are used to implement DEM reconstruction in a resolution-progressive manner. The first network aims at generating DEM at a resolution of 10 m/pixel, which is denoted as LDEMGAN-10 m. To train this network, the input data are downsampled LROC NAC images of 10 m/pixel and normalized SLDEM of 60 m/pixel (upsampled to the same grid spacing). Ground truth is downsampled LROC NAC DEMs at grid spacing

of 10 m/pixel. The second network intends to generate DEM of 2 m/pixel and is denoted as LDEMGAN-2 m. To train this network, the input data are LROC NAC image with resolution of 2 m/pixel, together with the LROC NAC DEM downsampled to 10 m/pixel, yet with similar pixel size to the NAC image. The ground truth is the LROC NAC DEM of its original resolution. For clearer demonstration, the 10 m reconstruction network and the 2 m reconstruction network are denoted as LDEMGAN-10 m and LDEMGAN-2 m, respectively. From this, it can be seen that the areas that satisfy the conditions for producing training data should include LROC NAC images, LROC NAC DEMs, and SLDEM at the same time.

According to the availability of data by the above requests, eight regions were selected for training the reconstruction networks. Some image samples of training data are shown in Figure 3: samples with different lunar surface features (Figure 3a), samples with different lighting conditions in the same area (Figure 3b), and samples after rotation or flipping (Figure 3c).

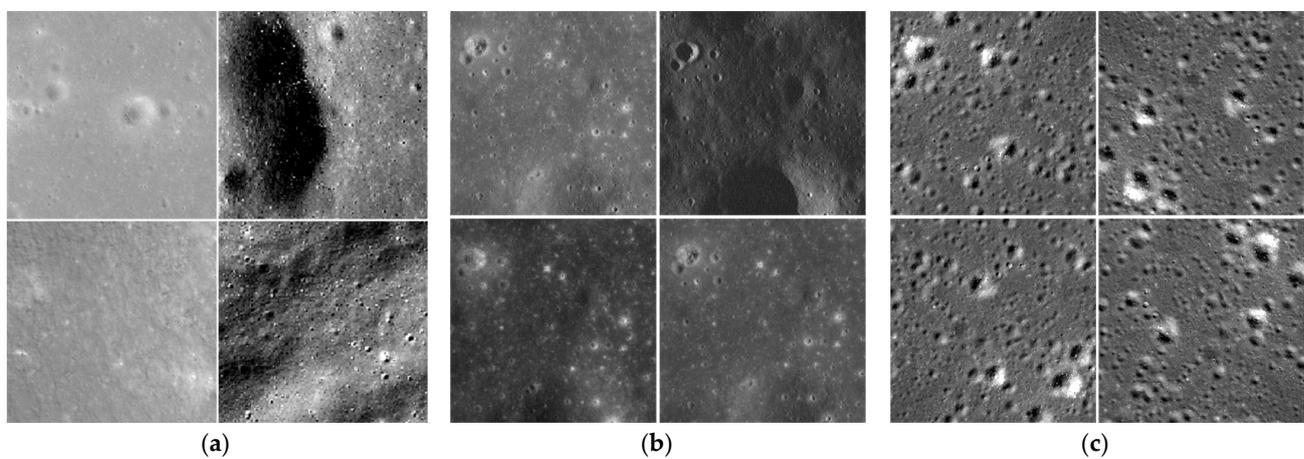


Figure 3. Examples of LROC NAC images from our training data: (a) images with different surface topography and features; (b) images with different lighting conditions in the same area; (c) images after rotation or flipping to simulate different sun azimuths.

A total of 89 scenes of LROC NAC images (2 m/pixel) and their corresponding LROC NAC DEM (2 m/pixel), together with SLDEM (60 m/pixel) covering these areas were selected to make training samples. A total of 6428 sets of training samples were made for training the first 10 m/pixel reconstruction network, and 4024 sets of training samples were made for training the second 2 m/pixel reconstruction network.

For LDEMGAN-10 m, the number of training epochs is 200, and the initial learning rate is 2×10^{-4} , which is linearly reduced to 10^{-6} after 100 epochs. The coefficients in Equation (11) are set to $\alpha = 0.5$, $\beta = 0.5$, and $\gamma = 0.8$, respectively. Next, LDEMGAN-2 m is obtained through migration learning based on LDEMGAN-10 m. In addition, during the LDEMGAN-2 m training, γ is reduced to 0.3 in Equation (11), and training epochs is set to 100. This experiment used the pytorch1.10 platform to build the network and was carried out on two RTX 3090 graphics cards. The training took a total of 5.7 h.

4.2. Testing

4.2.1. Testing Dataset

In order to test the reconstruction network, another four areas (including 25 LROC NAC images) different from training data areas containing LROC NAC DEM that aims at evaluating the reconstruction results are selected as test areas. The four areas are Apollo 11, Imbrium, Lichtenberg, and Hponds. The LROC NAC images of these areas are superimposed on SLDEM (Figure 4). The experimental areas are named according to what the LROC NAC DEMs contain.

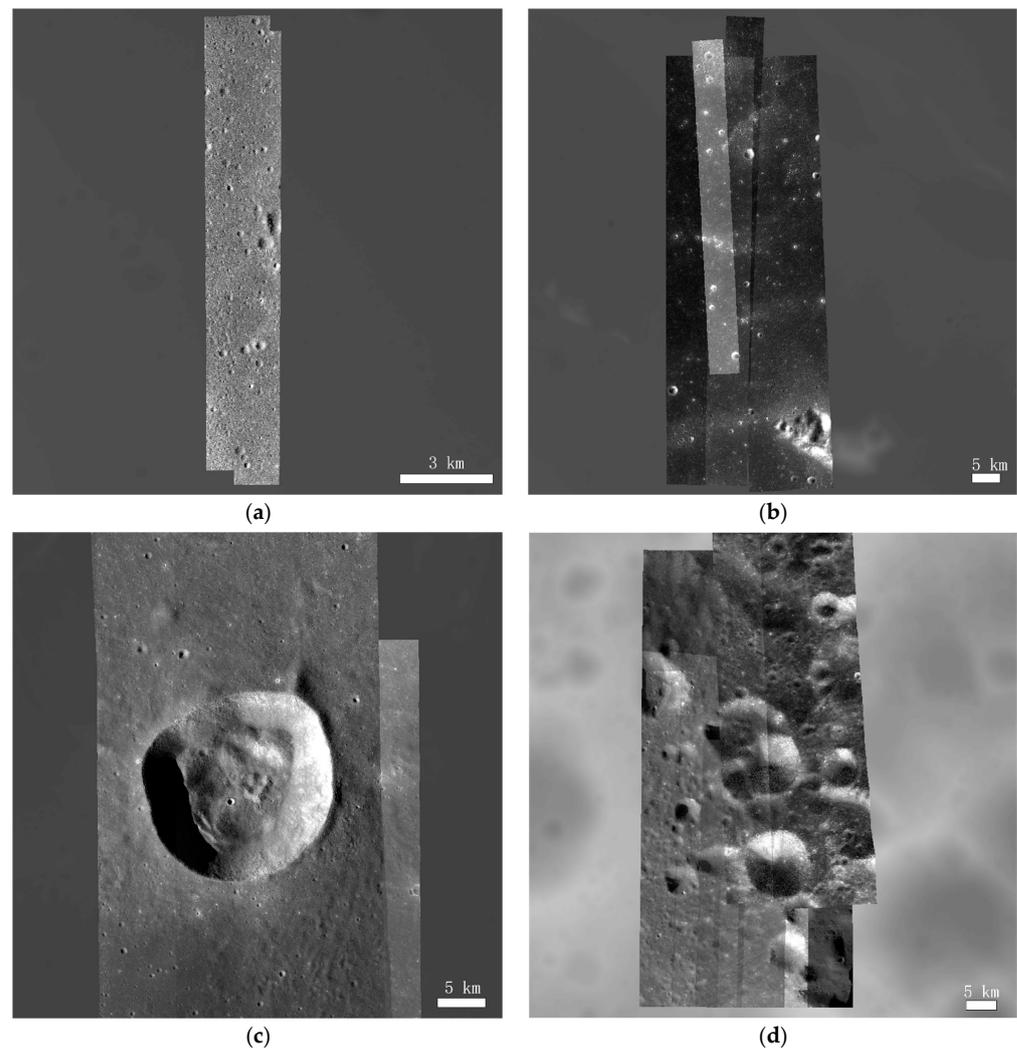


Figure 4. The LROC NAC images of the four test areas displayed on SLDEM: (a) Apollo 11; (b) Imbrium; (c) Lichtenberg; (d) Hponds.

Details of the four areas are shown in Table 1. The images used in each area are as follows: (1) Apollo 11: m150361817le, m150361817re, m150368601re; (2) Imbrium: m183697099le, m183704246re, m1106080949le, m1106080949re, m1175588941le, m1226179201re, m1265076277re; (3) Lichtenberg: m104955555le, m1096922142re, m1096929289le, m1096929289re, m1096936436le, m1096943583le, m1096936436re; (4) Hponds: m1100101542le, m1100101542re, m1102444668le, m1102444668re, m1110719713le, m1110719713re, m1113084806le, m1113084806re. The above data are downloaded from <https://ode.rsl.wustl.edu/moon/index.aspx> (accessed on 10 July 2022).

Table 1. Information of the four test areas.

ID	Test Area	Centre Latitude, Longitude (N, E/°)	Number of LROC NAC Images	SLDEM Resolution (m/pixel)
1	Apollo 11	0.77, 23.44	3	55.22
2	Imbrium	28.66, 333.88	7	55.51
3	Lichtenberg	31.68, 292.43	7	59.22
4	Hponds	41.79, 167.36	8	59.20

4.2.2. Reconstruction Results and Analysis

The reconstruction results of the four areas are shown in Figure 5. We use the LROC NAC DEMs of the experimental areas to evaluate the elevation accuracy of the reconstruction results. We choose two metrics, mean absolute error (MAE) and root mean square error (RMSE), to evaluate the reconstruction results. The two metrics can reflect the overall reconstruction performance of the statistics area, and the analysis of the local reconstruction effect is further analyzed by the profile graph and the absolute difference plot.

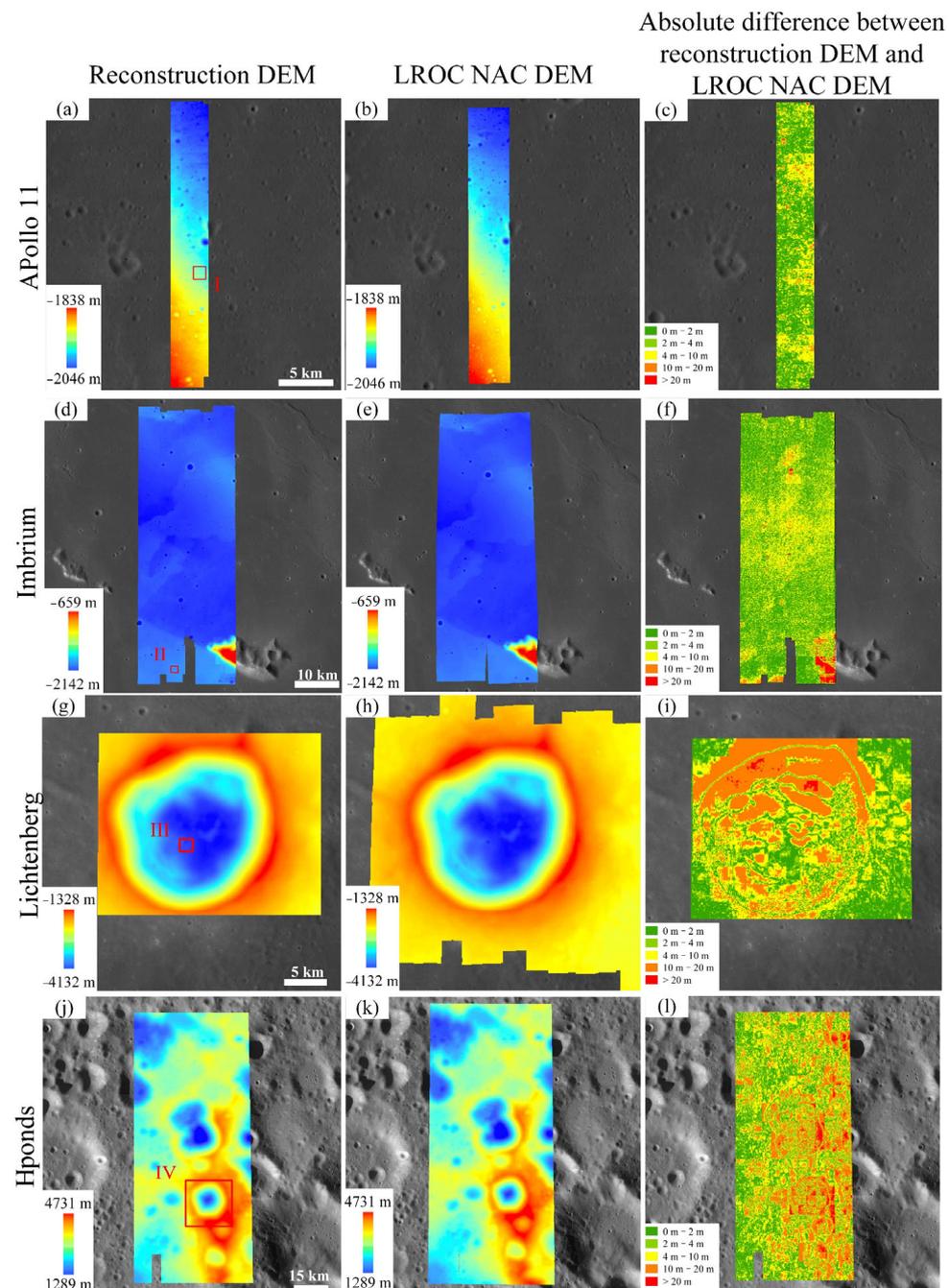


Figure 5. Reconstruction results and comparisons with LROC NAC DEMs. Each row shows one of the four experimental areas. The first column shows reconstruction results; the second column shows LROC NAC DEM as ground truth; the third column shows differences between the reconstruction results and LROC NAC DEMs.

The reconstruction effect of Apollo 11 (Figure 5a–c) is better than others in most parts, because the training data contain a considerable number of areas that have similar geomorphic characteristics and this area is relatively flat with less elevation variation. Approximately 50% of pixels have elevation differences less than 2 m, and more than 90% of pixels have elevation differences less than 10 m. The MAE of the whole experimental area is 2.76 m and the RMSE is 3.02 m, revealing high-precision reconstruction performance.

The reconstruction effect of Imbrium (Figure 5d–f) is similar to that of Apollo 11, and high reconstruction accuracy has been achieved in most parts of the area, the MAE and the RMSE is less than 4 m for the whole area, and more than 85% of pixels have elevation differences less than 10 m. The region with relatively large error in the Figure 5f bottom right corner is Mons La hire, which is a hillside with large topographic relief. In this region, the variation in image brightness (change in gray scale) is less than the other regions, yet the elevation change is relatively large; therefore, the reconstruction network may not be able to extract much effective information, leading to a comparatively large reconstruction error.

Experimental area Lichtenberg (Figure 5g–i) contains a large impact crater with diameter of 19 km and depth of 2.2 km. As illustrated in Figure 5i, large reconstruction errors mainly occur in the area with large topographic relief and less gray scale variation, such as near the upper edge of the impact crater.

The Hponds experimental area (Figure 5j–l) is located in the Chandler area on the far side of the Moon. Impact craters are densely and overlappingly distributed in this area. There are a large number of impact craters in this area with diameter more than 500 m and depth more than 100 m, which brings the existence of shading regions with brightness close to 0 in the impact crater. Both the large variation in elevation and shadows caused larger absolute reconstruction errors than the other test areas.

According to the above analysis of the reconstruction results, there are no obvious artifacts in the reconstruction results. The reconstructed elevation accuracy is related to the elevation and brightness ranges of the input area. The statistics of the reconstruction errors in the four areas are listed in Table 2. As can be found in Table 2, experimental areas Apollo 11 and Imbrium, which have moderate terrain fluctuation, have lower proportions of elevation errors (i.e., the differences between reconstructed results and LROC NAC DEMs) than the other two, which have relatively drastic topographic reliefs. The MAE and RMSE in these four areas also reflect this trend.

Table 2. The overall statistics of the reconstruction errors in the four areas.

Area	Reconstruction Error <2 m (%)	Reconstruction Error <4 m (%)	Reconstruction Error <10 m (%)	MAE (m)	RMSE (m)
Apollo 11	48.32	73.06	91.54	2.76	3.02
Imbrium	42.27	70.47	85.67	3.03	3.64
Lichtenberg	27.21	45.11	68.89	9.26	10.52
Hponds	25.52	45.95	62.51	9.43	10.65

4.2.3. Detailed Analysis of the Reconstruction Results in Subareas

In order to obtain a more detailed analysis of the reconstruction results, four typical subareas (numbered I–IV in Figure 5a,d,g,j) are selected to illustrate more elaborate analysis. Figure 6 shows the zoomed subareas, the corresponding SLDEM, the reconstruction results obtained by LDEMGAN, LROC NAC DEMs as ground truth for comparisons, and sectional views of the reconstruction results and gray scale profiles of the LROC NAC image, respectively. From Figure 6, it can be seen that the reconstructed DEMs from the monocular LROC NAC images have good consistency with LROC NAC DEMs, showing much finer elevation information than SLDEM, which provides control information.

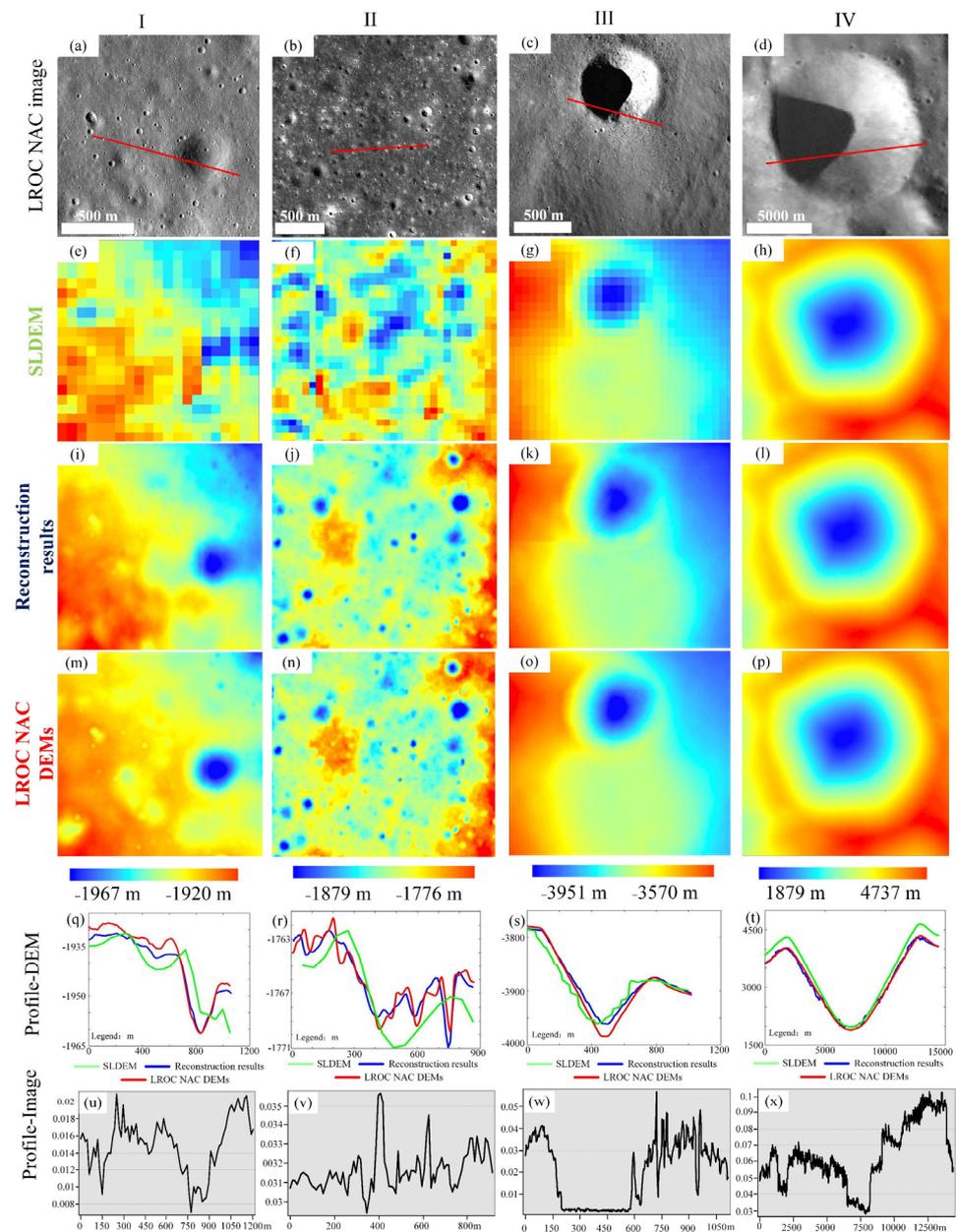


Figure 6. More detailed analysis of the results in the four subareas I–IV annotated in Figure 5. Each column shows one subarea. Row 1: LROC NAC image of 2 m/pixel showing the zoomed subareas. Row 2: SLDEM of 60 m/pixel. Row 3: network reconstruction results of 2 m/pixel. Row 4: LROC NAC DEM of 2 m/pixel as ground truth for comparison purposes. Row 5: DEM profiles of the four subareas with their profile locations shown in the first row. The green curves represent SLDEM, the blue curves represent network reconstruction results, and the red curves represent LROC NAC DEMs. Row 6: gray scale profiles of the LROC NAC image.

In subarea I (Figure 6a,e,i,m), the MAE of the reconstruction result is 1.26 m and the RMSE is 1.84 m, as demonstrated in Table 3. The maximum elevation error is 6.43 m, which is relatively small compared to the other subareas. From the elevation profiles (Figure 6d), it is shown that LDEMGEN can recover detailed DEM that coincides with the ground truth. Even the control information provided by SLDEM is with coarse grid and has shift in the bottoms of the curves, which demonstrate that the details in the image have more impact on the result.

Table 3. Reconstruction error of subareas.

Subarea	Degree of Elevation Variation (m)	Reconstruction Error <2 m (%)	MAE (m)	RMSE (m)	Max Error (m)
I	47.68	88.53	1.26	1.84	6.43
II	103.55	80.10	1.93	2.26	8.20
III	381.20	65.51	7.28	8.63	19.70
IV	2858.62	38.72	9.43	10.33	28.82

In subarea II (Figure 6b,f,j,n), when SLDEM is greatly different from the LROC NAC DEMs, as shown in Figure 6f,n, the reconstruction result is still highly consistent with the ground truth, from which it can be deduced that the input LROC NAC image takes more credit on recovering depth values. The MAE of the reconstruction result in this subarea is 1.93 m, the RMSE is 2.26 m, and the maximum elevation error is 8.20 m.

In subarea III (Figure 6c,g,k,o), the reconstruction result shows consistency with SLDEM (Figure 6g), yet with more details, with a reconstruction MAE of 7.28 m, RMSE of 8.63 m, and a maximum error of 19.70 m compared to the ground truth LROC NAC DEM (Figure 6k). It can be seen from the profile of this subarea (Figure 6s) that the center location of the crater in the reconstruction result coincides more with the ground truth than that of SLDEM, but there is an elevation drift of the reconstruction result from the real bottom value in the shadow region of the crater as the gray scale profile shown in Figure 6w. We found that when image brightness of the region is close to 0, the reconstruction network would be unable to translate much effective information from the image and may rely more on SLDEM to predict the elevation.

As can be seen from Figure 6a–c, the scales of subareas I, II and III are similar. The reconstruction errors rise with the increase of the degree of the elevation variation in these subareas.

In subarea IV (Figure 6d,h,p), the reconstruction result of a large impact crater with diameter of approximately 13.2 km is shown. From the profile (Figure 6t), the reconstruction results are better than the SLDEM (Figure 6h), especially on the edges of the craters. However, the RMSE is relatively larger than other subareas, which is 10.33 m, and the maximum error is 28.82 m. This is because the data are normalized at an imagery size of 256×256 to relative scale values before inputting into LDEMGAN and are recovered to absolute values according to the elevation range of this area. In consequence, deep craters or terrain with large elevation change would result in a relatively large absolute error.

In summary, based on the reconstruction results of these four subareas, the LDEMGAN can produce DEM with rich detail comparable to the NAC DEMs with the assistance of control information supplied by SLDEM. The accuracy of the reconstruction results is related to both the brightness variation of the input image and the degree of regional topographic relief. When the input images have less shading (the gray scale value is close to 0) and the degree of elevation variation is less than 100 m, the percentage of the reconstruction error less than 2 m (less than the input image grid spacing) can be more than 80%, the MAE can be less than 2 m, and the RMSE can reach below 2 m.

4.3. Discussion

4.3.1. Effect of Control Information

As demonstrated in the previous sections, the reconstruction results may rely more on the input image. In order to verify the necessity of control information instead of considering the process as just a style transfer operation from image to elevation, the reconstruction results with and without control information are compared in this subsection. The Apollo 11 landing area is selected as the test area, the input images are LROC NAC images (m150361817re, m150361817le), and the input low-resolution DEM for providing control information is downsampled LROC NAC DEMs (nac_dtm_apollo11) with grid spacing of 10 m/pixel. Here, using downsampled LROC NAC DEM instead of SLDEM can verify the improvement of the reconstruction effect under ideal conditions such that the

low-resolution data are completely correct. At the same time, considering that the degree of elevation change mentioned in Section 4.2 will affect the absolute error, the result of accuracy evaluation in this subsection applies relative elevation without scale recovery (i.e., the value range is $[0, 1]$) so that the error magnitude of each input datum is the same.

Figure 7 shows the comparison results using different inputs: using only LROC NAC image as input or using both LROC NAC image and coarse DEM as input. For better visualization, the relative elevation values are stretched to $[0, 255]$ for display in Figure 7. Figure 7c,h,m shows the reconstruction result taking only LROC NAC images as input. Although the trend of the elevation changes of these reconstruction results are consistent with the real lunar surface, which can indicate the locations of big craters, the extent of elevation change is inaccurate, especially in the areas that are relatively flat. In this case, the MAE of predicated results using only image is 0.11, the RMSE is 0.18, and the maximum mean error is 0.21. After adding the downsampled low-resolution DEM (coarse DEM) as input (Figure 7d,i,n), the reconstruction of elevation change extent is greatly improved while maintaining the detail. The error of each pixel of the reconstruction results is less than 0.14, the MAE is less than 2×10^{-3} and the RMSE of the pixels of all predicated results is 4.2×10^{-3} , which indicates great improvement in the reconstruction. This experiment verifies the necessity of using coarse DEM as control information, indicating that by adding coarse DEM in the reconstruction network, the low-frequency part can be preserved and noise can be reduced through the reconstruction process.

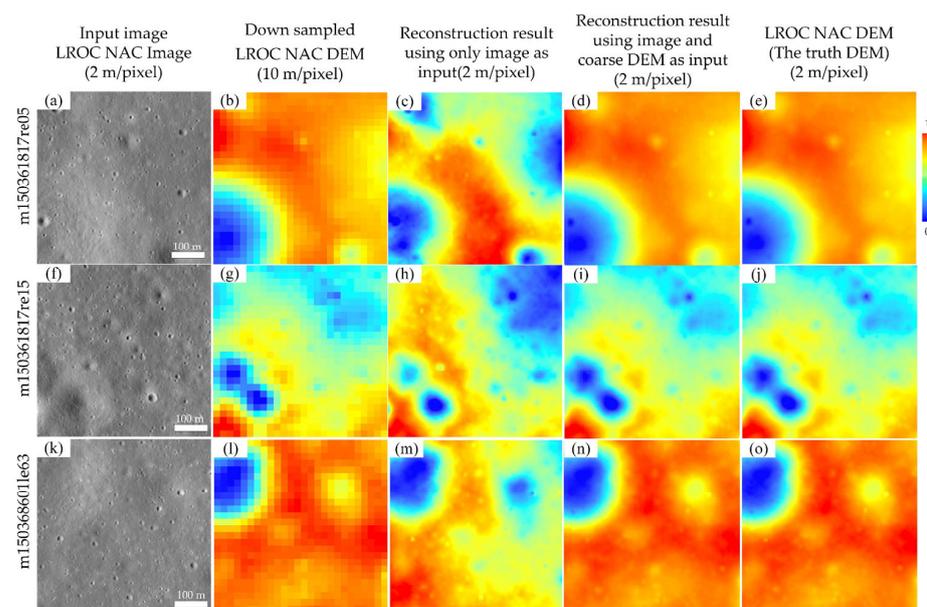


Figure 7. Comparisons of image reconstruction results with and without the help of coarse DEM. Each row shows results of one area. Column 1: The input image of LROC NAC image of 2 m/pixel. Column 2: The coarse DEM obtained by downsampling LROC NAC DEM from 2 m/pixel to 10 m/pixel. Column 3: The reconstructed result taking only image as input. Column 4: The reconstructed result using both image and coarse DEM as input. Column 5: The ground truth DEM of LROC NAC DEM (2 m/pixel).

4.3.2. Influence of Light Conditions

Similar to the traditional SFS method, the foundation of recovering elevation information from images through LDEMGEN is image photometric content changes, such as diverse intensity values due to reflection of lights from the topographic relief of the lunar surface. Under different lighting conditions, such as solar altitude angle, the imaged landforms of the lunar surface vary a lot, especially in the regions with severe elevation changes. Although during the training process, images taken under multiple solar altitude angles are considered to build the training dataset, relatively extreme cases, such as images

captured at too high or too low solar altitude angles, are excluded to ensure network convergence. Therefore, in this subsection, elevation reconstructions are carried out on the images captured under different lighting conditions. The tested images also include some with solar altitude angles that are not included during reconstruction network training. In this way, how different illumination conditions affect reconstruction results can be analyzed.

We selected 29 images captured at different solar altitude angles (see Table 4 for details), combined their SLDEM covering the same area as input, and reconstructed DEMs through LDEMGAN established in Section 4.1, and part of the test results are shown in Figure 8. On the whole, the reconstruction results (Figure 8e–h) of each image captured at different solar altitude angles (Figure 8a–d) are visually consistent, and the maximum RMSE of all reconstructed results is 3.27 m, which shows the network can adapt to the change in lighting conditions to maintain the consistency of reconstruction results.

Table 4. Solar altitude angle information of the experimental images used for testing influence of light conditions.

ID	Image	Solar Altitude Angle (°)	ID	Image	Solar Altitude Angle (°)
1	m1171733209le	0.22	16	m1159956344re	43.01
2	m1277607697le	5.04	17	m1129325970re	44.44
3	m102014464re	8.24	18	m1114014396re	46.40
4	m102000149re	10.25	19	m1114007294re	47.39
5	m1126986303re	15.50	20	m1121074527re	48.91
6	m1157600009re	16.36	21	m1121081627re	49.94
7	m1123426954re	21.27	22	m1197661005le	58.69
8	m1169345323re	23.45	23	m1282310415le	60.55
9	m1123448288le	24.38	24	m106719774le	62.77
10	m135032851le	28.86	25	m1192915976re	71.81
11	m113799518re	35.79	26	m1118716779re	76.14
12	m104362199re	36.03	27	m1116358477le	76.28
13	m129133239re	36.87	28	m1149338748re	76.90
14	m1129347305re	41.49	29	m1254093845le	80.50
15	m1129340193re	42.47			

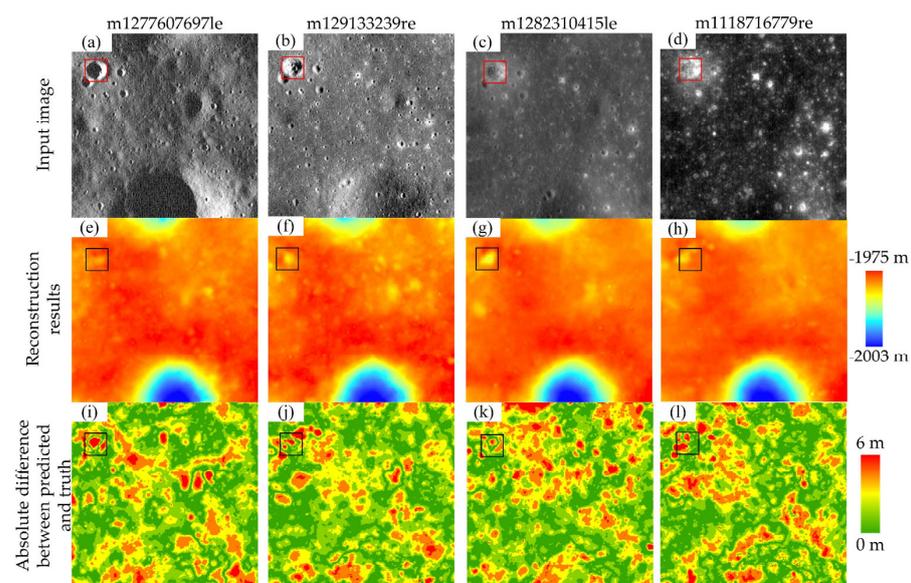


Figure 8. Comparison of reconstruction results using images captured at different solar altitude angles. Each column shows one result of one image. Row 1: the input image of LROC NAC image, solar altitude angle: m1277607697le, 5.04°; m129133239re, 36.87°; m1282310415le, 60.55°; m118716779re, 76.14°. Row 2: the reconstruction results. Row 3: the difference between the reconstruction results (2 m/pixel) and the LROC NAC DEM (ground truth, 2 m/pixel).

On the local scale, when the SLDEM due to the resolution limitation is also unable to provide effective information of the region to network, too high or too low solar altitude angle can lead to a decline in reconstruction accuracy. Especially, impact craters that cannot be shown in SLDEM (the region marked in red in image and in black in reconstruction results) have a noticeable difference. When the solar altitude angle is lower than 25° (as shown in Figure 8a, m1277607671e, solar altitude angle 5.04°), the RMSE of the marked region is 5.88 m. When the image brightness (gray scale value) of the region is close to 0, the reconstruction network could be unable to translate much effective information from the image. This results in a large reconstruction error in this region (the RMSE is 5.88 m). When the solar altitude angle is higher than 70° (as shown in Figure 8d, m1118716779re, solar altitude angle 76.14°), the RMSE of the marked region is 6.23 m, and the RMSE of the whole image is 4.48 m. The high solar altitude angle causes overexposure of the image. In this kind of image, small impact craters in the image are bright entirely or mostly scattered on the edge without distinguishable intensity changes, which leads to a certain degree of increase in reconstruction error.

According to the experiment results shown and analyzed in this subsection, it is proved that the network can be applied to different lighting conditions through the training method described in Section 4.1, and the different elevation restoration results caused by diverse lighting conditions have limited impact on the reconstruction results with acceptable errors.

4.3.3. Discussion of Experimental Results

In terms of network structure, the generator network of our proposed LDEMGAN adopts the U-Net structure, which takes both image and coarse DEM as input. The coarse DEM, providing low-frequency information in the network, is taken as control information for depth reconstruction (this is confirmed by the experiments in Section 4.3.1). However, this is based on the assumption that the input image and DEM are accurate and their spatial positions are consistent, which makes the accuracy of the reconstruction result affected by the quality and spatial consistency of the input data. In other words, the offset and systematic error in the spatial position of the input data will lead to unsatisfactory reconstruction results. In this work, as registration error is less than one grid spacing of the coarse DEM, the network is able to handle this error level, yet more research is needed to be done concerning precise registration and systematic error elimination in future work.

The input of LDEMGAN is a single image and a low-resolution DEM, which is not limited by stereo image coverage. The LDEMGAN is less affected by illumination changes (see Section 4.3.2 for details). In Section 4.2, the LROC NAC DEM as the ground truth was generated by stereophotogrammetry, and the RMSE and MAE in Table 2 are able to demonstrate that the DEM reconstructed by LDEMGAN is very close to the results reconstructed by stereophotogrammetry. Therefore, LDEMGAN is able to be used for DEM reconstruction of a wider area of the Moon compared to stereophotogrammetry. In addition, it is worth noting that the evaluation indicators MAE and RMSE used in this work are aimed at accuracy in a statistical sense, which can reflect overall reconstruction accuracy, but cannot reflect the spatial distribution of the error. In this paper, the spatial distribution of errors is mainly analyzed by the visual effect of the absolute difference plot (third column of Figure 5) and the profile of the subregion (fifth row of Figure 6). Because no obvious regularities affecting the reconstruction accuracy were found from the absolute difference maps or profile maps, the in-depth analysis was not continued in this paper. If the subsequent application requires an in-depth study of the anisotropy or spatial distribution law of the reconstruction effect, more relevant indicators will be needed for experiments and evaluation.

From Figure 6, it can be seen that the reconstruction results of the network in the area with large changes in elevation are not as accurate as that with small changes. We suppose that this may be related to the following reasons. (1) The elevation variation of the region where the training data was made is small, resulting in less ability of the

network to learn to extract such informative features. (2) The scale recovery is carried out according to Equation (13). When two areas with identical mean value but different standard deviations (sd), e.g., greater degree of elevation change results in larger sd value, the one with larger sd value probably leads to larger errors according to this scale recovery method. (3) Scale recovery only depends on low-resolution DEM. The low-resolution DEM may have systematic and random errors that are biased from the real lunar surface. These errors will reflect on the reconstruction results through scale recovery. From the above discussion, in future work, we need to firstly improve the normalization and scale recovery mechanisms, and then add various training data so that the network may identify and learn the different degree of elevation change. In this case, systematic errors caused by data quality may be detected and corrected for better DEM reconstruction performance.

In subarea IV shown in Figure 6t, it seems that the reconstruction results are not affected by the errors of SLDEM, and the reconstruction results are close to the ground truth. The mean value of SLDEM in this subarea is close to the ground truth, but its standard deviation is large. We deduce that gradient loss L_{grad} included in the loss function in Equation (11) of LDEMGAN may take this credit of eliminating errors. When the network finds that the gradient information extracted from the image is inconsistent with the elevation change of SLDEM, it will trust the high-frequency information obtained from the image more.

5. Conclusions

In this paper, according to the current situation that there are high-resolution images and low-resolution DEMs covering the whole lunar surface, but few high-resolution stereo pairs, we propose a high-resolution lunar DEM reconstruction method based on GAN. This method can reconstruct a high-resolution DEM using a single high-resolution image and a much lower-resolution DEM in the corresponding area. The grid spacing of the reconstructed DEM is as high as that of the high-resolution input image.

The experimental results verify the feasibility and effectiveness of the proposed method and verify the necessity of using much coarser DEM taken as control information in the network, instead of only using it to recover absolute depth. The RMSE of the reconstruction results can reach about 2 m in the area where the degree of elevation variation is less than 100 m, and the RMSE value ranges from around 3 m to 10 m without considering the degree of the elevation variation in large area reconstruction. In addition, the reconstruction accuracy is proved stable under nonextreme illumination conditions (the solar altitude angle is between 25° and 75°). In the future, we will continue to conduct research on anisotropy of the reconstruction results, so as to further optimize and improve the network.

Due to the current and near-future data situation on the lunar surface, the proposed approach has great potential for building global high-resolution DEMs for the lunar surface to support more scientific and engineering applications, such as topographic characterization, landing site selection, navigation, positioning, path planning, decision support, etc.

Author Contributions: Conceptualization, K.D., Y.W. and Y.L.; methodology, Y.L. and Y.W.; software, Y.L. and Y.W.; validation, M.P., W.W. and Z.L.; formal analysis, Y.W.; investigation, K.D., Y.W. and M.P.; resources, W.W. and Z.L.; writing—original draft preparation, Y.L.; writing—review and editing, K.D. and Y.W.; project administration, Y.W.; funding acquisition, K.D., Y.W. and M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key Research and Development Program of China (grant 2018YFB1305004) and the National Natural Science Foundation of China (grants 41941003 and 41771488).

Data Availability Statement: Publicly available datasets were analyzed in this study. The datasets can be found here: <https://ode.rsl.wustl.edu/moon/>.

Acknowledgments: The authors thank all those who worked on the NASA Planetary Data System archive to make the LROC imagery and SLDEM available.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kokhanov, A.A.; Karachevtseva, I.P.; Zubarev, A.; Patraty, V.; Rodionova, Z.; Oberst, J. Mapping of potential lunar landing areas using LRO and SELENE data. *Planet. Space Sci.* **2018**, *162*, 179–189. [[CrossRef](#)]
2. Liu, Z.Q.; Di, K.C.; Peng, M.; Wan, W.; Liu, B.; Li, L.; Yu, T.; Wang, B.; Zhou, J.; Chen, H. High precision landing site mapping and rover localization for Chang'e-3 mission. *Sci. China-Phys. Mech. Astron.* **2015**, *58*, 1–11. [[CrossRef](#)]
3. Liu, Z.Q.; Di, K.C.; Li, J.; Xie, J.; Cui, X.; Xi, L.; Wan, W.; Peng, M.; Liu, B.; Wang, Y.; et al. Landing site topographic mapping and rover localization for Chang'e-4 mission. *Sci. China Inf. Sci.* **2020**, *63*, 140901. [[CrossRef](#)]
4. Wu, B.; Li, F.; Ye, L.; Qiao, S.; Huang, J.; Wu, X.; Zhang, H. Topographic modeling and analysis of the landing site of Chang'E-3 on the Moon. *Earth Planet. Sci. Lett.* **2014**, *405*, 257–273. [[CrossRef](#)]
5. Wu, B.; Li, F.; Hu, H.; Zhao, Y.; Wang, Y.; Xiao, P.; Li, Y.; Liu, W.C.; Chen, L.; Ge, X.; et al. Topographic and Geomorphological Mapping and Analysis of the Chang'E-4 Landing Site on the Far Side of the Moon. *Photogramm. Eng. Remote Sens.* **2020**, *86*, 247–258. [[CrossRef](#)]
6. Di, K.; Oberst, J.; Karachevtseva, I.; Wu, B. Topographic Mapping of The Moon in the 21st Century: From Hectometer to Millimeter Scales. *ISPRS J. Photogramm. Remote Sens.* **2020**, *XLIII-B3-2020*, 1117–1124. [[CrossRef](#)]
7. Smith, D.E.; Zuber, M.T.; Jackson, G.B.; Cavanaugh, J.F.; Neumann, G.; Riris, H.; Sun, X.; Zellar, R.S.; Coltharp, C.; Connelly, J.; et al. The Lunar Orbiter Laser Altimeter Investigation on the Lunar Reconnaissance Orbiter Mission. *Space Sci. Rev.* **2010**, *150*, 209–241. [[CrossRef](#)]
8. Haruyama, J.; Hara, S.; Hioki, K.; Iwasaki, A.; Morota, T.; Ohtake, M.; Matsunaga, T.; Araki, H.; Matsumoto, K.; Ishihara, Y.; et al. Lunar Global Digital Terrain Model Dataset Produced from SELENE (Kaguya) Terrain Camera Stereo Observations. In Proceedings of the Lunar & Planetary Science Conference, Woodlands, TX, USA, 19–23 March 2012.
9. Li, C.L.; Ren, X.; Liu, J.J.; Wang, F.F.; Wang, W.R.; Yan, W.; Zhang, G.H. A New Global and High Resolution Topographic Map Product of the Moon from Chang'e-2 Image Data. In Proceedings of the Lunar & Planetary Science Conference, Woodlands, TX, USA, 16–20 March 2015.
10. Barker, M.K.; Mazarico, E.M.; Neumann, G.; Zuber, M.; Haruyama, J.; Smith, D. A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera. *Icarus* **2016**, *273*, 346–355. [[CrossRef](#)]
11. Wu, B.; Liu, W.C.; Grumpe, A.; Wöhler, C. Construction of pixel-level resolution DEMs from monocular images by shape and albedo from shading constrained with low-resolution DEM. *J. Photogramm. Remote Sens.* **2018**, *140*, 3–19. [[CrossRef](#)]
12. Henriksen, M.R.; Manheim, M.R.; Burns, K.; Seymour, P.; Speyerer, E.; Deran, A.; Boyd, A.; Howington-Kraus, E.; Rosiek, M.; Archinal, B.; et al. Extracting accurate and precise topography from LROC narrow angle camera stereo observations. *Icarus* **2017**, *283*, 122–137. [[CrossRef](#)]
13. Kirk, R.L.; Barrett, J.M.; Soderblom, L.A. Commission IV, Working Group IV/9. Photoclinometry Made Simple . . . ? In Proceedings of the “Advances in Planetary Mapping 2003” Workshop, Houston, TX, USA, 22 March 2003.
14. Lohse, V.; Heipke, C. Multi-image shape-from-shading: Derivation of planetary digital terrain models using clementine images. *Remote Sens.* **2004**, *35*, 828–833.
15. Grumpe, A.; Belkhir, F.; Wöhler, C. Construction of lunar DEMs based on reflectance modelling. *Adv. Space Res.* **2014**, *53*, 1735–1767. [[CrossRef](#)]
16. Horn, B.K. Understanding Image Intensities. *Artif. Intell. Rev.* **1977**, *8*, 201–231. [[CrossRef](#)]
17. Zhang, R.; Tsai, P.S.; Cryer, J.; Shah, M. Shape from shading: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 690–706. [[CrossRef](#)]
18. Barron, J.T.; Malik, J. High-Frequency Shape and Albedo from Shading using Natural Image Statistics. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011.
19. Danzl, R.; Scherer, S. Integrating Shape from Shading and Shape from Stereo for Variable Reflectance Surface Reconstruction from SEM Images. In Proceedings of the 26th Workshop of the Austrian Association for Pattern Recognition, Graz, Austria, 10 September 2002; pp. 281–288.
20. Chen, Z.Y.; Wu, B.; Liu, W. Mars3DNet: CNN-Based High-Resolution 3D Reconstruction of the Martian Surface from Single Images. *Remote Sens.* **2021**, *13*, 839. [[CrossRef](#)]
21. Khan, F.; Salahuddin, S.; Javidnia, H. Deep Learning-Based Monocular Depth Estimation Methods-A State-of-the-Art Review. *Sensors* **2020**, *20*, 2272. [[CrossRef](#)] [[PubMed](#)]
22. Eigen, D.; Puhrsch, C.; Fergus, R. Depth Map Prediction from a Single Image using a Multi-Scale Deep Network. In Proceedings of the 28th Conference on Neural Information Processing Systems (NIPS), Montreal, QC, Canada, 8–13 December 2014.
23. Eigen, D.; Fergus, R. Predicting Depth, Surface Normals and Semantic Labels with a Common Multi-Scale Convolutional Architecture. In Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 2650–2658.
24. Laina, I.; Ruppel, C.; Belagiannis, V.; Tombari, F.; Navab, N. Deeper Depth Prediction with Fully Convolutional Residual Networks. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 239–248.

25. Ma, X.; Geng, Z.; Bie, Z. Depth Estimation from Single Image Using CNN-Residual Network. Semantic Scholar. Available online: <http://cs231n.stanford.edu/reports/2017/pdfs/203.pdf> (accessed on 21 July 2021).
26. Fu, H.; Gong, M.M.; Wang, C.; Batmanghelich, K.; Tao, D. Deep Ordinal Regression Network for Monocular Depth Estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2002–2011.
27. Facil, J.M.; Ummenhofer, B.; Zhou, H.; Montesano, L.; Brox, T.; Civera, J. CAM-Convs: Camera-Aware Multi-Scale Convolutions for Single-View Depth. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 11826–11835.
28. Wolk, D.; Ma, F.C.; Yang, T.-J.; Karaman, S.; Sze, V. FastDepth: Fast Monocular Depth Estimation on Embedded Systems. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 6101–6108.
29. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Commun. AcM.* **2020**, *63*, 139–144. [[CrossRef](#)]
30. Jung, H.; Kim, Y.; Min, D.; Oh, C.; Sohn, K. Depth prediction from a single image with conditional adversarial networks. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 1717–1721.
31. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Springer: Cham, Switzerland, 2015; Volume 9351, pp. 234–241.
32. Li, J.; Jia, J.J.; Xu, D. Unsupervised Representation Learning of Image-Based Plant Disease with Deep Convolutional Generative Adversarial Networks. In Proceedings of the Chinese Control Conference (Ccc), Wuhan, China, 25–27 July 2018; pp. 9159–9163.
33. Lore, K.G.; Reddy, K.; Giering, M.; Bernal, E.A. Generative adversarial networks for depth map estimation from RGB video. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018.
34. Tao, Y.; Muller, J.P.; Conway, S.J.; Xiong, S. Large Area High-Resolution 3D Mapping of Oxia Planum: The Landing Site for the ExoMars Rosalind Franklin Rover. *Remote Sens.* **2021**, *13*, 3270. [[CrossRef](#)]
35. Tao, Y.; Muller, J.P.; Xiong, S.; Conway, S.J. MADNet 2.0: Pixel-Scale Topography Retrieval from Single-View Orbital Imagery of Mars Using Deep Learning. *Remote Sens.* **2021**, *13*, 4220. [[CrossRef](#)]
36. Hao, C.; Xuanyu, H.; Oberst, J. Pixel-Resolution DTM Generation for The Lunar Surface Based on a Combined Deep Learning and Shape-From-Shading (SFS) Approach. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Nice, France, 6–11 June 2022; Volume V-3-2022, pp. 511–516.