IMAGE-TO-DEM TRANSLATION WITH CONDITIONAL ADVERSARIAL NETWORKS OF DEPTH ESTIMATION BASED ON MONOCULAR IMAGES ON THE MOON. K. Ibuka[1], K. Ogino[2], M. Goto[2], M. Ohtake[3] and H. Demura[3] [1]School of Computer Sci. and Eng., Univ. of Aizu, (koichiro.ibuka@gmail.com), [2] Graduate School of Computer Sci. and Eng., Univ. of Aizu, [3]Aizu Research Center for Space Informatics (ARC-Space), Univ. of Aizu, Aizu-Wakamatsu City, Fukushima 965-8580, JAPAN (demura@u-aizu.ac.jp).

Introduction: Digital Elevation Models (DEMs) of the lunar surface are popular for planning lunar explorations and for studies on topography and geology [1]. We need high-resolution DEMs because LRO NAC DEMs with the highest resolution (~2 m/pixel) show little coverage of the lunar area with only 470 pieces [2]. The little coverage is derived from the small number of suitable data for stereogrammetry as a method primarily used for DEMs generation [3]. Time consuming process for making DEMs manually also prevents from generating larger coverage of DEMs [4].

The aim of this study is to generate high-resolution DEMs automatically based on monocular images. We propose a method to generate DEMs from visible images by using Pix2Pix [5], a model of Generative Adversarial Networks (GAN). In related work, Liu et al. (2022) [6] proposes a method to obtain high-resolution DEMs generated by a GAN and inputting them into its reconstruction network. In this study, we propose a method to generate high-resolution DEMs by GAN only. Our generated DEMs are evaluated by Mean Average Error (MAE), Root Square Mean Error (RMSE), and Line Profile against the existing DEMs. Two goals of this study are to compare the accuracy of the generated DEMs with the results of Liu et al. (2022) [6] and to evaluate its validity of the DEM by GAN only.

Methods: We propose a method to generate DEMs from lunar visible images using Pix2Pix. The architecture of the Pix2Pix is shown in Figure 1. It consists of two modules: the image generator ***G*** and the image discriminator ***D***. In this study, ***G*** generates fake DEMs and ***D*** determines whether they are generated by ***G*** or exist as data products. ***G*** generates highly accurate DEMs by repeating this process so that ***G*** deceive ***D***.

This study has two phases in training with Pix2Pix. We use pairs of low-resolution DEM and visible image to learn large size of terrain features. Next, we train Pix2Pix with hyper-parameters derived from the first phase to capture fine-grained ones using pairs of high-resolution DEM and visible image.

We evaluate the accuracy of the generated DEMs using the three metrics. RMSE and MAE indicate similarity between the real DEM and the generated one overall. Two Line profiles of them also gives similarity.

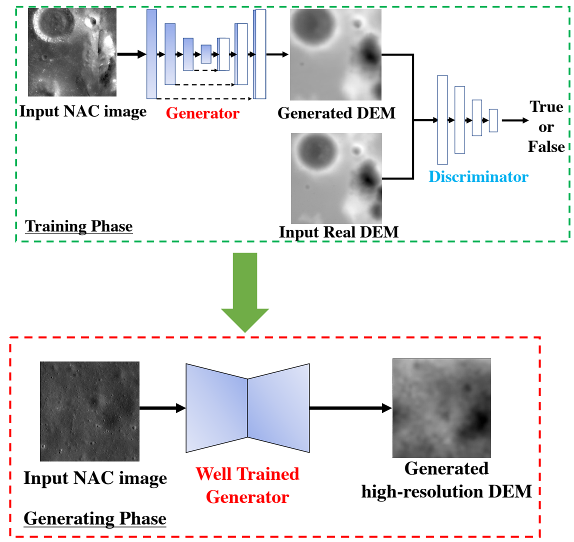

Figure 1: DEM Generation Flow

Dataset: We use SLDEM2013[7], SLDEM2015[8], and LROC NAC orthophoto images [9] as our datasets. This is because these DEMs are abundance covering region within ±60° latitudes. SLDEM2013 is a data product generated using topographic data obtained from the high-resolution (10m/pixel) stereo Terrain Camera (TC) onboard SELENE (KAGUYA). SLDEM2015 is corrected for 2013 with data measured by the Lunar Orbiter Laser Altimeter (LOLA) onboard the Lunar Reconnaissance Orbiter (LRO) obtained through 2015. LROC NAC Orthophoto images are stereo-rectified and ortho-rectified image taken from two cameras on board the Lunar Reconnaissance Orbiter.

The training dataset consists of two different pairs; pairs of SLDEM2013 and LROC NAC orthophoto images, and pairs of SLDEM2015 and LROC NAC orthophoto images. LROC NAC Orthophoto images are downsampled to match the resolution of SLDEM2013 and SLDEM2015. The misalignment between the NAC image and the DEM is corrected by Affine transformation based on the mutual information. Each image is cropped at 320×320 pixel and then cropped again to 256×256 pixel after applying random rotation and random vertical and horizontal flip.

Training settings: We prepared 830 low-resolution pairs, and 8430 high-resolution pairs. We trained a network on low-resolution and high-resolution each pair at 5000 epochs and with a batch size of 8.

Results: We evaluate the generated DEMs in the region of "highland ponds" (latitude 39.04˚ - 44.53˚, longitude 165.21˚ - 169.52˚). Table 1 shows a result of RMSE and MAE by comparison between this study and Liu et al (2022) [6]. Both errors in this study show six times larger than those of them.

Figure 2 shows a line-profile of the generated DEMs with the smallest errors. The generated DEM reconstructs the change of regional topography although the change of local one deviates from the LROC NAC DEM.

Figure 3 shows another line-profile of the generated DEMs with largest errors. This figure displays failure of generated DEM. Possible causes should be investigated.

Table1: Comparison of this research with LDEMGAN

	MAE (m)	RMSE (m)
This Research	62.11	61.24
LDEMGAN [6]	9.43	10.65

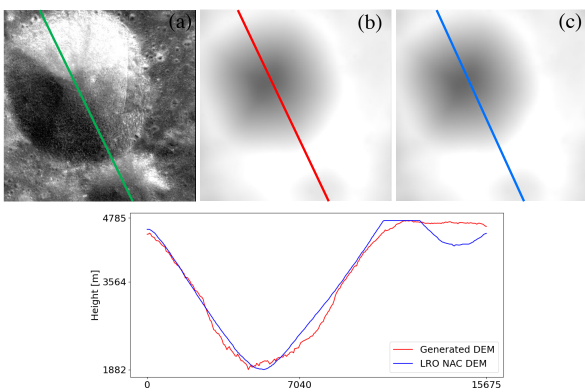

Figure 2: A Good example of generated DEM

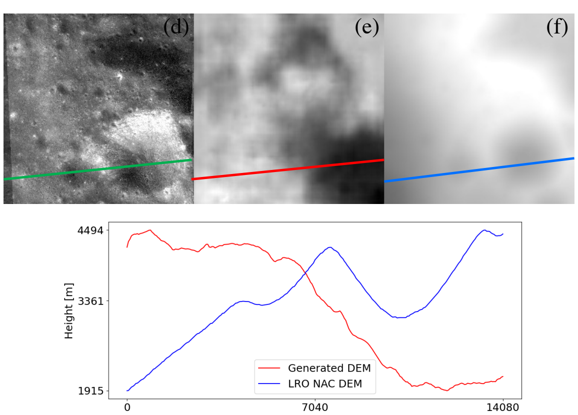

Figure 3 An example of failure in DEM reconstruction

Discussion: The accuracy of the DEMs generated in this study is lower than in previous studies. We interpret this low accuracy as quality of the training dataset. This would be caused by low accuracy of registration between visible images and DEM. A few pixels of misalignment seem to prevent from reconstructing fine features.

Illumination conditions of images in the dataset are considered to affect quality of DEMs. The different conditions bring various appearances with shadings and shadows. We use LROC NAC images with incidence angle within 35˚ - 65˚. Narrowing the range of incidence angles and training the model for specific conditions is expected to improve the accuracy of DEM.

Conclusion: This study compares the validity of generated DEMs from lunar images with previous study by Liu et al. (2022) [6]. The error of RMSE and MAE in this study is 6 times as large as that of LDEMGAN [6]. Improvement in the accuracy of DEMs is expected by revising correction of misalignments among data products, and by considering unification of the illumination conditions.

Acknowledgments: This research was supported by the Sasakawa Scientific Research Grant from The Japan Science Society. We are pleased to acknowledge the considerable assistance of Profs. Honda, Hirata, Okudaira, and Ogawa. We thanks to all those who make and published LROC NAC images, DEMs, SLDEM2013 and 2015.

References: [1] Baker, M. K. *et al.* (2016) *Icarus, 273, 346-355*. [2] Henriksen, M. R. *et al.* (2020) *Lunar Surface Science Workshop,* LPI Contribution 2241, id. 5084. [3] Chen, H. *et al.* (2022) *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.,* V-3-2022, 511-516. [4] Onodera, K. *et al.* (2020) *JAXA-RR,19-006*. [5] Isola, P. *et al.* (2017) *CVPR*, 1125-1134. [6] Liu *et al.* (2022) *Remote Sens., 14, 21, 5420-5440.* [7]https://data.darts.isas.jaxa.jp/pub/pds3/sln-l-tc-5-sldem2013-v1.0/
[8]http://imbrium.mit.edu/EXTRAS/SLDEM2015/
[9] https://wms.lroc.asu.edu/lroc/rdr_product_select