Application of Denoising CNN for Noise Suppression and Weak Signal Extraction of Lunar Penetrating Radar Data

Haoqiu Zhou 1,2, Xuan Feng 1,2,3,*, Zejun Dong 1,2, Cai Liu 1,3 and Wenjing Liang 1,2

Abstract: As one of the main payloads mounted on the Yutu-2 rover of Chang’E-4 probe, lunar penetrating radar (LPR) aims to map the subsurface structure in the Von Kármán crater. The field LPR data are generally masked by clutters and noises of large quantities. To solve the noise interference, dozens of filtering methods have been applied to LPR data. However, these methods have their limitations, so noise suppression is still a tough issue worth studying. In this article, the denoising convolutional neural network (CNN) framework is applied to the noise suppression and weak signal extraction of 500 MHz LPR data. The results verify that the low-frequency clutters embedded in the LPR data mainly came from the instrument system of the Yutu rover. Besides, compared with the classic band-pass filter and the mean filter, the CNN filter has better performance when dealing with noise interference and weak signal extraction; compared with Kirchhoff migration, it can provide original high-quality radargram with diffraction information. Based on the high-quality radargram provided by the CNN filter, the subsurface sandwich structure is revealed and the weak signals from three sub-layers within the paleo-regolith are extracted.

Keywords: Chang’E-4; lunar penetrating radar (LPR); convolutional neural network (CNN); noise suppression; weak signal extraction

1. Introduction

After China’s Chang’E-3 (CE-3) mission was completed, the Chang’E-4 (CE-4) probe was launched on 8 December 2018, and achieved the first-ever soft landing on the farside of the moon on January 3, 2019. The landing site of CE-4 is on the floor of Von Kármán crater in the South Pole-Aitken (SPA) basin [1]. With a diameter of approximately 2500 km and the deepest depth of 13 km, the SPA basin is the largest impact structure on the moon and can provide important clues to the internal structure of the moon [2–7]. Recent works indicate that the near-surface materials measured by the CE-4 rover were mainly the ejecta from the Finsen crater, with possible contributions from the Alder, the Von Kármán L, and the Von Kármán’ L’ craters [7–17]. It is also convincing that the ejecta from the Finsen crater probably contains the lunar mantle materials [18–21].

The Yutu-2 rover, carried by CE-4, is mounted with a lunar penetrating radar (LPR) which is used to detect the lunar subsurface structure [22]. Similar to CE-3 [23,24], the LPR possesses two channels (CH-1 and CH-2) with center frequencies of 60 and 500 MHz, respectively [23–29]. The CH-2 consists of one transmitting antenna and two receiving antennas (also known as CH-2A and CH-2B) of different offsets [30–32]. The mission of CH-2 is to map the fine structure of the lunar regolith [23–32].

The CH-2 LPR amplitude-modulated and frequency-modulated signals are always masked by clutters and noises of large quantity [33]. The clutters with relatively low
frequency may generate from the coupling of antennas and the instrument system, as there are many kinds of payloads mounted on the rover [22]. The processing of LPR data is challenging especially in noise suppression. In previous researches, many filtering methods have been applied to LPR data, such as band-pass filtering [34], bi-dimensional empirical mode decomposition (BDEMD) [35], complete ensemble empirical mode decomposition (CEEMD) [36], mathematical morphology filtering [37], and variational mode decomposition (VMD) [38]. However, these methods have their limitations and cannot provide a satisfactory result. More importantly, due to the attenuation of electromagnetic waves, the signals of great depth are quite weak and immersed within the noises. The noises and weak signals are of similar amplitudes and frequencies and cannot be easily classified using traditional denoising methods. Therefore, it is of vital importance to develop a more effective filtering method for LPR data processing. Convolutional neural network (CNN) [39–41] is a type of deep learning network architecture that is widely applied to various fields. In 2017, Zhang et al. proposed a CNN model for image denoising [42,43]. The denoising CNN model innovatively removed the pooling layers from the initial CNN model, as pooling layers may erase the characteristics of high-frequency noise.

In this article, the denoising CNN framework is applied to CE-4 LPR data filtering. The low-frequency clutters used in the CNN training are extracted from CE-3 LPR data using the empirical mode decomposition (EMD) [44]. Therefore, the achievement of noise suppression can also verify that the low-frequency clutters embedded in the CE-4 are the same as that of CE-3 and are generated from the instrument system of the rover. Based on the well trained CNN filter, the low-frequency clutters and high-frequency noises are both suppressed in the field LPR profile and the weak signals in the deep are extracted. To illustrate the superiority of the CNN filter on noise suppression, a classic band-pass filter and a mean filter are respectively conducted as comparisons. For discussing the results of weak signal extraction, the combined methods of the band-pass filter and the mean filter as well as the combined methods of the band-pass filter and the Kirchhoff migration are respectively conducted as comparisons. Finally, based on the high-quality radargram provided by the CNN filter, the subsurface sandwich structure is revealed and the weak signals from three sub-layers within the paleo-regolith are extracted.

2. Materials and Methods

2.1. The Initial CH-2B LPR Data

The walking route for the Yutu-2 rover on the moon for the first 12 lunar days is shown in Figure 1a [11]. The curve is ~320 meters in length. A series of processing steps, including the traces amending, the traces selecting, the time lag adjustment, the attenuation compensation, and the background removal [10,11,25–31] (The details are in Table A1) are conducted to obtain the initial CH-2B profile. The original profile and its average frequency amplitude spectrum are respectively shown in Figure 1b,c. Considering that the center frequency of the CH-2B LPR signal is 500 MHz with a bandwidth of 450 MHz [23,24], the radargram in Figure 1b is masked by low-frequency clutters and high-frequency noises of large quantity. Simultaneously, as the signals from the deep are very weak and their amplitudes are similar to the noise, the signals below 400 ns are more severely affected by the noise interference and the reflections from the strata are invisible or hard to recognize. Therefore, the key point of this article is to suppress the noises and extract the deep weak signals.
Figure 1. (a) The traveling route of Yutu-2 rover within the first 12 lunar days [11]. (b) The initial CH-2B LPR radargram acquired along the path in (a) without filtering. (c) The average frequency amplitude spectrum of the radargram in (b), which is masked by low-frequency clutter and high-frequency noise of large quantities.

2.2. Methodology of Denoising CNN

The key idea of deep learning is to deepen the depth of the network for enhancing the accuracy of nonlinear matching. The CNN is a method of deep learning. It integrates the input data into a hierarchical stack of operations such as convolution operations, convergent operations, and nonlinear activation function mapping, and extracts features from the original data [39–41]. This process is called “feed-forward operation.” The last layer of the CNN compares the operation result with the target task, converts it into the objective function, obtains the error between the calculated value and the real value, and then uses the error back-propagation (BP) algorithm to propagate the error back. The error is assigned to the neurons of each layer, and then the parameters of each layer can be updated. This process is called “feed-back operation.” The iterations will continue until the objective function converges, which means the network is well trained. When new data are input, they can be predicted by the network’s feed-forward operation.

Denoising CNN is a CNN framework for image denoising. It innovatively removed the pooling layer from the initial CNN model. The framework of denoising CNN is shown in Figure 2 [42,43]. In this model, the batch normalization (BN) and the rectified linear unit (Relu) are applied. These two training techniques are used to avoid the vanishing gradient problem [42,43]. The mathematical form of BN is given by [42,43]:

\[ P_{out} = \frac{P_{in} - \text{mean}(P_{in})}{\sqrt{\text{var}(P_{in})}} \cdot \gamma + \beta, \]  

(1)
where $P_{\text{in}}$ and $P_{\text{out}}$ represent the input and output of a layer. Mean($P_{\text{in}}$) and var($P_{\text{in}}$) are the average value and standard deviation of $P_{\text{in}}$. $\gamma$ and $\beta$ are the learnable parameters.

The Relu is a type of nonlinear activation function widely applied in deep learning [42,43]. It always follows the convolutional layer, and its mathematical expression is shown as follow:

$$P_{\text{out}} = \max(0, P_{\text{in}} \cdot w + b),$$

where $w$ and $b$ are the convolutional kernel and the threshold, respectively.

### 3. Results

#### 3.1. CNN Training

The key of CNN training is to use the simulated GPR data with low-frequency clutters and high-frequency noises as the training data. The GPR simulations for different types of relative permittivity models are conducted using the finite-difference time-domain (FDTD) method. The models as well as their simulated data are shown in Figure 3a,b. The models contain different target combinations with different relative permittivity.

As we have discussed in Section 2.1, the LPR data are masked by low-frequency clutters and high-frequency noises of large quantities. The low-frequency clutter is considered...
to generate from the coupling of antennas and the instrument system, so it is relatively stable. As CE-4 is the backup of CE-3, we decide to extract 200 low-frequency clutters from CH-2B data of CE-3 using empirical mode decomposition (EMD) [44]. If the filtering result is satisfactory, the speculation for the source of low-frequency clutters can also be proved. The extracted clutters as well as their average frequency amplitude spectrum are respectively shown in Figure 4a, b. Besides, the high-frequency noise can be easily simulated by Gaussian noise. Subsequently, the low-frequency clutters and high-frequency noises are embedded into the simulated GPR data using the formula below:

\[ \text{Image} = \text{Label} + \alpha_1 LC + \alpha_2 HN, \]  

(3)

where \text{Label} represents the simulated GPR data; \text{LC} and \text{HN} are low-frequency clutters and high-frequency noises, respectively; \text{Image} denotes the data masked by \text{LC} and \text{HN}. The \( \alpha_1 \) and \( \alpha_2 \) are the intensity coefficients. The greater \( \alpha_1 \) and \( \alpha_2 \) correspond to the lower signal-to-noise ratio (SNR).

![Figure 4. The 200 extracted low-frequency clutters from Chang’E-3 LPR data and its average frequency amplitude spectrum. (a) The extracted low-frequency clutters. (b) Average frequency amplitude spectrum of the signals in (a).](image)

Figure 4. The 200 extracted low-frequency clutters from Chang’E-3 LPR data and its average frequency amplitude spectrum. (a) The extracted low-frequency clutters. (b) Average frequency amplitude spectrum of the signals in (a).

Compared with the random noises, the clutters are more stable and, more importantly, time-related. Therefore, before adding clutters, the lengths of training radar profiles are re-sized into the length of clutters, which can make the training data remaining the features of clutters. Besides, the additions of \text{LC} are conducted for a single trace of the simulated data at a time. For each trace, we randomly draw one of the 200 clutters and add it to the signal.

Here, a test is conducted to determine the ranges of these two parameters in the CNN training process. The thresholds should contain most possible cases of different noise levels and the maximums of these two parameters should guarantee that the signals are almost immersed in clutters or noises. Besides, wider ranges of \( \alpha_1 \) and \( \alpha_2 \) will result in generating more training images which will decelerate the convergence of CNN training and make it more difficult and time-consuming. Therefore, we need to select the smallest ranges of \( \alpha_1 \) and \( \alpha_2 \) while covering most of the possible noise cases.

The tests with \( \alpha_1 = 0, 1, 1.5, \) and 2 and \( \alpha_2 = 0, 1, 3, \) and 5 are shown in Figures 5 and 6, which are based on the simulated data of the first radargram in Figure 3b. It is evident that the signals from the target gradually become invisible with the increase in \( \alpha_1 \) and \( \alpha_2 \).
To provide quantitative results, the image entropy is conducted for different images with different $\alpha_1$ and $\alpha_2$. The image entropy is given by [45–47]:

$$R = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} a^2(m, n)}{\sum_{m=1}^{M} \sum_{n=1}^{N} a^4(m, n)},$$

(4)

where $M$ and $N$ denote the vertical and horizontal size of a radar profile; $a(m, n)$ represents the pixel value of the radargram at $(m, n)$ position. This parameter is known to be an easily computable, accurate approximation to entropy [45–47]. The larger $R$ value indicates the greater interference of clutters and noises.

Figure 5. The comparison of radar profiles with different parameter $\alpha_1$. (a) The normalized radargram with $\alpha_1 = 0$. (b) $\alpha_1 = 1$. (c) $\alpha_1 = 1.5$. (d) $\alpha_1 = 2$.

Figure 6. The comparison of radar profiles with different parameter $\alpha_2$. (a) The normalized radargram with $\alpha_2 = 0$. (b) $\alpha_2 = 1$. (c) $\alpha_2 = 3$. (d) $\alpha_2 = 5$.

Figure 7a shows the image entropy of the image with different $\alpha_1$ and $\alpha_2$; the blue and orange curves represent the clutters added and random noise added cases, respectively. Figure 7b presents the derivatives of the curves in Figure 7a. Both curves of image entropy rise with the increase in $\alpha_1$ and $\alpha_2$; when $\alpha_1 = 2$ and $\alpha_2 = 5$, the curves gradually become flat corresponding to the zero values of derivatives in Figure 7b. In this case, the clutters and noises become the main ingredients of the images. It is also feasible to keep increasing the $\alpha_1$ and $\alpha_2$, but the images will become complete clutters and noises without signals. These useless images will not strengthen the training result but make the training more time-consuming. Therefore, $\alpha_1 \in [0, 2]$, $\alpha_2 \in [0, 5]$ are reasonable thresholds that are the smallest and can cover most of the possible noise cases.
The image entropy analysis of Figure 7. (a) The derivatives of the curves in (a).

(b) The normalized radar-profiles of the image entropy with different $\alpha_1$ and $\alpha_2$.

To expand the training data set and enhance the performance of CNN, the widely used data augmentation method [42,43] is subsequently applied. The augmentation contains horizontal and vertical flip. Finally, the augmented data are applied to the CNN training. The convergence curve of the loss function is shown in Figure 8. The objective function converges at approximately 200 iterations.

![Convergence curve of the loss function.](image)

Figure 8. Convergence curve of the loss function.

3.2. Model Denoising Test

In this section, the integrated model of lunar regolith [27,35] is applied to test the CNN filter. This model is built based on the actual environment of lunar regolith and stochastic media [27,35]. The relative permittivity model is shown in Figure 9. The top layer is a vacuum; the middle layer is the lunar regolith; the bottom layer is the bedrock. The grain sizes ascend with the increasing depth, due to the different weathering degree. The simulated GPR data, the noise embedded data, and the CNN filtering result are shown in Figure 10. The CNN filter can suppress both the low-frequency clutters and the high-frequency noise well. The filtered image is very close to the original image, the hyperbolas are well restored. Therefore, the results convince that the CNN for the GPR data denoising is feasible and the CNN model has been well trained. Besides, the signals in the radargram with both low-frequency clutter and high-frequency noises in Figure 10c are almost totally immersed in the interference and are hard to recognize. In the test, the amplitudes of the clutters and noises are set to be close to those of the signals, the case of which is similar to the field LPR data with the great depths. However, the CNN filter shows its stable ability to suppress the interference and restore the signals.
Figure 9. The integrated relative permittivity model of lunar regolith [27,35].

Figure 10. Model denoising test. (a) Simulated data; (b) data with low-frequency clutters; (c) data with low-frequency clutters and high-frequency noises; (d) denoising result; (e) comparison of average frequency amplitude spectrum.

3.3. Denoising and Weak Signal Extraction of LPR Data

In this section, the well-trained denoising CNN model is applied to the denoising and weak signal extraction of LPR data. To well illustrate the superiority of the denoising CNN method, the classic band-pass filter is first conducted as the comparison method. The results of the band-pass filtering are shown in Figure 11a–c. The radargram contains much less noise compared to the initial radargram in Figure 1c which is also evident in the comparison of the average frequency amplitude spectra in Figure 11c. However, the signals below 400 ns are still masked with large quantities of noise and cannot be recognized. To explain the phenomenon, the S transformation [48] is conducted to analyze the property of the noise embedded in the LPR data after band-pass filtering. Four signals are extracted in Figure 11a and their S transformation results are presented in Figure 11d–g. It is obvious that the frequencies of the clutters and noises embedded below 400 ns are within the frequency band of LPR signals reflected from the subsurface structures and are also in the frequency band of the filter in Figure 11b. Therefore, the frequency-based filter such as the band-pass filter cannot suppress the clutter and noise below 400 ns.
This phenomenon can be interpreted to be the result of the magnification of the ringing effect and noise. Due to the attenuation of electromagnetic waves [29], the signals from the deep are weak and their amplitudes are close to the ringing signals. In order to make the deep signals more visible, the attenuation compensation [10,11,25–31] is necessary, resulting in the magnification of ringing signals and noise. This interference is of the same frequency band as the useful LPR signals and is difficult to suppress.

Subsequently, the CNN filter is applied to the LPR data. As a frequency-based filter such as the band-pass filter cannot suppress the clutter and noise below 400 ns, we simultaneously conducted the classical mean filter as a comparison method with the CNN filter. The mean filter is a typical linear filter, which involves a template to the target pixel on the image that includes the neighboring pixels around it (the 3 × 3 sized pixel region around the target pixel constitutes a filtering template). The filtering process removes the target pixel itself and replaces the original pixel value with the average of all the pixels in the template.

Figure 12a,b presents the results of the mean filter and CNN filter. Besides, Figure 12c,d shows the radargram after both band-pass and mean filtering and the radargram after Kirchhoff migration, respectively. For analyzing the processing results, Figure 12e compares the average frequency amplitude spectra of the results. The result after only mean filtering cannot suppress the low-frequency clutter and high-frequency noise, which is also evident in the spectra comparison. As the purple curve represents the mean filter, it is evident that the mean filter is not effective in suppressing the low-frequency clutter and high-frequency noise. The green curve represents the CNN filter, and it is evident that the CNN filter is effective in suppressing the low-frequency clutter and high-frequency noise. The blue curve represents the band-pass filter, and it is evident that the band-pass filter is effective in suppressing the low-frequency clutter and high-frequency noise. The red curve represents the Kirchhoff migration, and it is evident that the Kirchhoff migration is effective in suppressing the low-frequency clutter and high-frequency noise.

Figure 11. Band-pass filtering test. (a) The radargram after band-pass filtering. The signals marked by black lines are the signals used for the S transformation analysis. (b) Parameters of the band-pass filter. (c) The average frequency amplitude spectrum of the LPR data before and after band-pass filtering. (d–g) Present the S transformation results of the marked signals in (a).
both band-pass and Kirchhoff migration, respectively. For analyzing the processing results, Figure 12e compares the average frequency amplitude spectra of the results.

The result after only mean filtering cannot suppress the low-frequency clutter and high-frequency noise, which is also evident in the spectra comparison. As the purple curve in Figure 12e shows, the mean filter can suppress large quantities of high-frequency noises but the result is much worse than the CNN and band-pass filtering. Besides, the mean filtering has almost no attenuation ability for low-frequency clutters.

The CNN filtering can suppress the clutters and noises well, especially for the signals below 400 ns (Figure 12b). An evident interface at ~500 ns and another unclear interface at ~600 ns can be found, which are pointed out by the double-way black arrows (Figure 12b). Furthermore, we test a combined method of conducting both the band-pass and the mean filters (Figure 12c). The CNN and the combined methods are similar in the spectra results, but the image quality of the combined method result falls flat. The imaging result of the combined method is still masked with many clutters and noises and only the unclear interface at ~500 ns can be observed whereas the signals at ~600 ns are immersed in the noise. Some strips of interference (the blue arrows in Figure 12c), which can also be
observed in the original radargram (Figure 1a) or the mean filtering result (Figure 12a), are not completely removed. However, these interferences are well attenuated in the CNN filtering results. Therefore, the result of CNN filtering can present a radargram with higher quality compared with the combined method of band-pass and mean filtering.

Besides, the migration processing after band-pass filtering, which is commonly adopted in the previous studies [9–11], can present a similar result (Figure 12b) to that of the CNN filtering. However, the result of migration has attenuated the diffracted waves and is much more suitable for stratigraphic divisions [9–11] whereas the CNN filter can obtain an original radargram with the image of high quality and information of diffracted waves. The diffracted wave contains important information for property analysis of LPR data [30,32,34,37,38,49–52] and the high-quality image can guarantee the accuracy of the analysis. Overall, the result of migration can provide clues to stratigraphic divisions and the result of the CNN filter can provide high-quality radargram for property analysis.

3.4. The Subsurface Sandwich Structure at CE-4 Landing Site

In Section 3.3, the superiority of the CNN method has been verified. Based on the extracted signals from the radargram, a subsurface structure can be summarized as Figure 13 shows. The radargram shows a fine-coarse-fine sandwich structure [9–11,53].

![Figure 13. Stratigraphic divisions result and interpretation. The base of the figure is the radargram after CNN filtering. The vertical axis is the transformed depth from two-way traveling time with velocity assumed to 0.16 m/ns [11]. The radargram shows a fine-coarse-fine sandwich structure. The black curves indicate the interfaces of the sandwich structure. The dotted curves represent the interfaces of the sub-layers within the paleo-regolith.](image)

The top 0–10 m is the fine-grained regolith, which is mainly composed of unconsolidated materials accumulating gradually from the continuous impaction and gardening of the lunar surface materials [12]. The second layer with a depth of 10–35 m is the coarse-grained material which can be interpreted as the superposition of multiple ejecta sources from nearby craters [9–11]. The signals from this layer contain many reflections from subsurface boulders which have survived the impaction and gardening during the formation of the overlying fine-grained regolith.

The analysis for the signals >35 m is the focus of our studies. The third layer also contains fine-grained materials similar to the surface layer which can be interpreted to be the paleo-regolith. Besides, based on the CNN filtering result, three sub-layers can be observed within this layer. The interface at ~40 m is evident and continuous and another interface at ~48 m is unclear but also continuous. Although the depths are beyond the estimated penetrating depth [50], the CNN filter can also make the weak signal visible. The top sub-layer has a few reflections. This layer can be interpreted as a transition zone and may contain the mixture of the paleo-regolith material and the overlying coarse ejecta. The mixed material was formed during the placement of the overlying material. Furthermore,
the continuous reflected signals within the fine-grained material at ~40 m and ~48 m are not likely to be the reflections from the boulders within the layers but the signals from horizontal interfaces of two layers with different dielectric properties. Therefore, the paleo-regolith contains three kinds of materials with different dielectric properties and sources. The materials were probably excavated from more distant and different craters relative to the local impactions. However, it is still difficult to find multiple sources of different dielectric properties within such a small region. Considering the geologic content and the depths of the layers, a possible interpretation is that these materials may have different contents of FeO and TiO$_2$, which are related to the magmatic activities during the Imbrian period [7–10,14–17]. The material may be the ejecta containing the basalt material excavated from different nearby craters. Due to the different degrees of mixing with the local materials, the ejecta contains different contents of FeO and TiO$_2$, which results in the difference in dielectric properties. Besides, the formation time intervals of the three sub-layers are relatively longer than that of the first two layers (0–10 m and 10–35 m). Each sub-layer had been through long-time gardening before its overlying sub-layer was formed; this process of lunar surface homogenized the material and gradually formed the layered fine-grained paleo-regolith.

4. Discussions

4.1. Detailed and Quantitative Comparisons of Filtering Methods

In this section, we conducted detailed and quantitative comparisons of the methods mentioned in Section 3.3. The comparisons mainly focus on the denoising results of the LPR radargram below 35 m. Figure 14a–e presents the radar profiles below 450 ns (>~35 m) with no filtering, mean filtering, band-pass filtering, band-pass and mean filtering, and CNN filtering, respectively. The mean filtering and band-pass filtering have almost no effect on the suppression of noise and clutter interference. The blue arrows point out the strips of interference on the radargrams; this interference will increase the difficulties of stratigraphic interpretation. The combined method of band-pass and mean filtering and the CNN method are the two methods that can present the clear interface at 500 ns. However, the radargram provided by the combined method is not clean compared to the result of CNN filtering and some strips of interference are still visible.

Subsequently, three regions (the black rectangles in Figure 14d,e) on the results of the combined method and CNN method are extracted and presented in Figure 15. Figure 15a–c corresponds to the regions 1–3 in Figure 14d,e, respectively. The red up-toward and down-toward arrows indicate the signals at 500 and 600 ns, respectively. The results presented by the CNN filter are cleaner than that of the combined method. The events in radargram after CNN processing are more visible and continuous and can be well recognized. The signals on the results of the combined method, especially pointed by the down-toward arrows, are still masked by strong noise. Besides, the strips of interference pointed by the blue arrows are still evident on the results of the combined method but well suppressed on the CNN filtering results. Comparing the images in the rectangles, the CNN method reduces the effects of the distortion and well reconstructs and smooths the events.

Furthermore, an objective quantitative comparison is conducted based on the image entropy given by (4) in Section 3.1. The results are shown in Table 1. The image entropy of CNN filtering results are all lower than those of band-pass and mean filtering combined method. According to the analysis in Figures 5–7, Section 3.1, the low image entropy corresponds to low noise and clutter interference. Therefore, these results are consistent with the image analysis before and objectively verify the results.
Table 1. Image entropy comparison of the filtering results.

| Images               | CNN Filtering | Band-Pass and Mean Filtering |
|----------------------|---------------|-----------------------------|
| Figure 14d,e         | $1.9075 \times 10^6$ | $2.1871 \times 10^6$         |
| Figure 15a           | $5.7698 \times 10^4$  | $7.3160 \times 10^4$         |
| Figure 15b           | $6.3342 \times 10^4$  | $7.4902 \times 10^4$         |
| Figure 15c           | $5.5177 \times 10^4$  | $7.0545 \times 10^4$         |

Figure 14. Comparisons of radar profiles below 450 ns (depth below ~35 m). (a) Original radargram with no filtering processing. (b) Radargram after mean filtering. (c) Radargram after band-pass filtering. (d) Radargram after the combined method of band-pass and mean filtering. (e) Radargram after CNN filtering. The blue arrows indicate the strips of interference. The three rectangular regions 1–3 are magnified and presented in Figure 15.
4.2. Why the CNN Filter Performs Better?

Finally, we will discuss why the CNN filter performs better than the combined method of band-pass and mean filter. First, the spectra of the original data and the data after processing of the two methods are extracted from Figure 12e and presented in Figure 16. According to the spectra, the result of CNN filtering seems worse than the band-pass and mean filter; the curves of CNN result in the low-frequency and high-frequency region are farther from the horizontal axis compared with the combined method. We further calculate the mean integral relative errors (MIRE) between the original signals and the denoised results. The MIRE is given by:

\[ E = \frac{\int_{f} (a_{O}(f, x) - a_{D}(f, x)) df}{N_{x}}, \]

where \( a_{O} \) and \( a_{D} \) denote the original data and the denoised result in the \( f \)-\( x \) domain, respectively. \( f \) represents the frequency. \( N_{x} \) is the trace number. The \( E \) value can present the difference between the spectra of the original data and the denoised result. In most cases,
it is of a better denoising effect with a greater $E$ value. The comparison of the $E$ value is shown in Table 2. The results indicate that, according to the spectra and MIRE comparison, the combined method of band-pass and mean filter performs better than the CNN filter.

Figure 16. Detailed spectra comparison between the CNN filtering result and the combined band-pass and mean filtering result.

Table 2. Mean integral relative error comparison of the filtering results in Figure 12d,e.

| Filtering Method                      | Mean Integral Relative Error |
|---------------------------------------|------------------------------|
| CNN Filtering                        | $1.9278 \times 10^3$        |
| Band-Pass and Mean Filtering          | $2.4477 \times 10^3$        |

However, according to the image analysis and image entropy comparison, we have concluded that the CNN filter performs better. The results are contradictory but reasonable. The spectra analysis and the MIRE describe the features of the single trace signals but the image analysis and image entropy focus on the feature of the 2D images.

Essentially, a band-pass filter is a 1D filter that focuses on the spectra of single traces and cannot handle the matching of trace signals in the horizontal directions. Therefore, the band-pass filter cannot suppress the horizontal distortion such as the strips of interference in the radargram. By combining the mean filter, the result becomes better. However, the mean filter has limitations and cannot fully suppress the noise and clutter. According to the result in Figure 14b, the mean filter is sensitive to noise and should be performed after other filters like the band-pass filter. Besides, Figure 12e indicates that the mean filter has almost no attenuation ability for low frequency clutters which is because the effect of a mean filter depends on the size of a filtering template. Similarly, for the horizontal distortions with a low wave number such as the strips of interference, the mean filter will perform poorly.

The key of the CNN framework is the convolution layer which can be expressed as follows:

$$P_{in}(x, y) \ast w(x, y) = \sum_{x_0=a_1}^{a_2} \sum_{y_0=b_1}^{b_2} P_{in}(x_0, y_0)w(x - x_0, y - y_0),$$

where $P_{in}(x, y)$ and $w(x, y)$ are the input image and convolutional weight, respectively. $a_1$, $a_2$ and $b_1$, $b_2$ represent the horizontal and vertical ranges of the calculation region.

A CNN filter is essentially a 2D filter. The convolutional processing can extract the features of radargram and is performed for the whole 2D image in each layer. It can not only suppress the noise but also handle the matching of trace signals in the horizontal directions, which makes the events more evident, smoothing, and continuous in Figure 15. These advantages of CNN are the basis of the further geologic interpretations.

5. Conclusions

In this article, the denoising CNN framework is applied to the denoising of 500 MHz LPR data. The low-frequency clutter extracted from CE-3 LPR data using the EMD method
is applied to the training of CNN. The results verify that the low-frequency clutters embedded in the LPR data mainly came from the instrument system of the Yutu rover. Besides, the frequency-based filter such as the classic band-pass filter cannot suppress the interference due to the magnification of ringing clutters as the signals and clutters have the same frequency bands. Furthermore, compared with the combined methods of the band-pass filter and the mean filter, the CNN filter has better performance when dealing with the noise interference and weak signal extraction; compared with the combined methods of the band-pass filter and the Kirchhoff migration, it can provide original high-quality radargram with diffraction information. Based on the high-quality radargram provided by the CNN filter, the subsurface sandwich structure is revealed and the weak signals from three sub-layers within the paleo-regolith are extracted. The materials in the three sub-layers probably contain different contents of FeO and TiO$_2$, which are related to the magmatic activities during the Imbrian period. The materials may be the ejecta containing the basalt material from nearby craters.

**Author Contributions:** Conceptualization, H.Z., X.F. and Z.D.; data curation, H.Z. and Z.D.; formal analysis, H.Z., Z.D. and C.L.; funding acquisition, X.F. and W.L.; investigation, H.Z., X.F. and Z.D.; methodology, H.Z., X.F. and Z.D.; project administration, X.F. and W.L.; resources, C.L., X.F. and W.L.; software, H.Z.; supervision, C.L.; validation, H.Z., X.F. and Z.D.; visualization, H.Z., X.F. and Z.D.; writing—original draft, H.Z. and Z.D.; writing—review and editing, H.Z., X.F. and Z.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key Research and Development Program of China under Grant 2018YFC1503705, the National Natural Science Foundation of China under Grant 41974129, the special fund of Key Laboratory of Geo-Exploration Instruments, Ministry of Education under Grant 352190230537, and the Technology Development Program of Jilin Province under Grant 20180101091JC.

**Data Availability Statement:** The scientific data are available at Data Publishing and Information Service System of China’s Lunar Exploration Program (http://moon.bao.ac.cn (accessed on 19 January 2021)).

**Acknowledgments:** We thank the China National Space Administration for providing the scientific data used in this study. The scientific data are available at Data Publishing and Information Service System of China’s Lunar Exploration Program (http://moon.bao.ac.cn (accessed on 19 January 2021)).

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

**Appendix A**

The appendix contains a Table of data processing procedures. The procedures are conducted to obtain the initial radargram in Figure 1c.

**Table A1.** Data processing procedures before filtering.

| Processing                  | Details                                                                 |
|-----------------------------|-------------------------------------------------------------------------|
| Traces amending             | Adjusting the longitudinal displacement of traces, based on the phase of a strong reflection event. The rover might stop at some points on the way to collect other scientific data. But LPR never stops measurement, which resulted in the repeated acquisition of multiple traces at the same location. We average the repeated traces. |
| Traces selecting            | Conducting SEC makes deep information more visible. Removing the average background data. |
| Time lag adjustment         | There is a 28.203 ns delay for the start time of the transmitting antenna compared to the receiving antenna. |
| Attenuation compensation    |                                                                         |
| Background removal          |                                                                         |
