Article

Coupling Complementary Strategy to U-Net Based Convolution Neural Network for Detecting Lunar Impact Craters

Yuqing Mao ^{1,2,†}, Rongao Yuan ^{1,†}, Wei Li ¹ and Yijing Liu ^{1,*}

¹ School of Computer, Sichuan University, Chengdu 610065, China; maoyuqing@stu.scu.edu.cn (Y.M.); rgyuan@stu.scy.edu.cn (R.Y.); 2020223045114@stu.scu.edu.cn (W.L.)

² School of Cyber Science and Engineering, Sichuan University, Chengdu 610065, China

* Correspondence: liuyijing@scu.edu.cn

† These authors contributed equally to this work.

Abstract: Lunar crater detection plays an important role in lunar exploration, while machine learning (ML) exhibits promising advantages in the field. However, previous ML works almost all used a single type of lunar map, such as an elevation map (DEM) or orthographic projection map (WAC), to extract crater features; the two types of images have individual limitations on reflecting the crater features, which lead to insufficient feature information, in turn influencing the detection performance. To address this limitation, we, in this work, propose feature complementary of the two types of images and accordingly explore an advanced dual-path convolutional neural network (Dual-Path) based on a U-NET structure to effectively conduct feature integration. Dual-Path consists of a contracting path, bridging path, and expanding path. The contracting path separately extracts features from DEM and WAC images by means of two independent input branches, while the bridging layer integrates the two types of features by 1×1 convolution. Finally, the expanding path, coupled with the attention mechanism, further learns and optimizes the feature information. In addition, a special deep convolution block with a residual module is introduced to avoid network degradation and gradient disappearance. The ablation experiment and the comparison of four competitive models only using DEM features confirm that the feature complementary can effectively improve the detection performance and speed. Our model is further verified by different regions of the whole moon, exhibiting high robustness and potential in practical applications.

Keywords: convolutional neural network; deep learning; crater detection; moon

Citation: Mao, Y.; Yuan, R.; Li, W.; Liu, Y. Coupling Complementary Strategy to U-Net Based Convolution Neural Network for Detecting Lunar Impact Craters. *Remote Sens.* **2022**, *14*, 661. <https://doi.org/10.3390/rs14030661>

Academic Editors: Pawel Rotter, Wojciech Chmiel and Sławomir Mikrut

Received: 24 December 2021

Accepted: 27 January 2022

Published: 29 January 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

The moon is the first choice for human astronomical activities and space exploration activities, which are of great significance to human development. Impact craters are the most obvious and main morphological features on the lunar surface, which could provide important clues for studying the evolutionary history of the moon and space exploration. Thus, many efforts have been devoted to recognizing lunar impact craters, including artificial recognition [1–4], image transformation and segmentation [5–7], geoscience information analysis [8,9], and machine learning [10–12]. Artificial recognition is a method in which experts or other astronomers use telescopes to take pictures and mark impact craters manually in lunar images. However, with the growth of planetary data, manual extraction is too time-consuming and laborious. Feature matching uses typical image features, such as the annular structure of lunar craters as bases. Then, the crater is extracted by means of segmentation or edge fitting from these bases. Its precision is dependent on the feature manually selected, thus limiting its adaptability. The methods based on the image transformation and segmentation use different filtering and detection algorithms to recognize the image features of the lunar surface, while the method based on geoscience information analysis is to use the information from slopes, textures, and curvature of slopes

to gain insight into the impact crater. The two methods are susceptible to the complexity of the geographic environment and have limitations on the craters, with degradation on the edges and overlapping impact craters.

With the development of artificial intelligence, machine learning, as its core technique, possesses a strong learning capacity and can capture the useful information underlying complex data. Thus, it attracts increasing interest from various fields, including lunar impact crater detection. Some traditional machine learning methods, such as the support vector machine and decision fusion methods, were used to construct a classification model for the orthographic projection and elevation map data [10–12]. Despite some successes, the traditional machine learning methods generally rely on handcrafting features, which is also time consuming (DL) and prone to bias [13]. Compared to traditional machine learning, deep learning is more powerful in capturing complex relationships and can avoid hand-selected features. In fact, planetary data tend to be massive; thus, in principle, deep learning is more suitable to identify large moon images. Jin et al. [14] used Fast R-CNN [15] to detect the impact crater in high-resolution orthographic projection images. Although its accuracy and recall were reported to be 92.96% and 89.19%, respectively, such high accuracy was only limited at the landing site of Chang'E 4, rather than the whole moon; thus, its generality to other terrains needs to be validated. Ali-Dib et al. [16] used the weak supervised deep learning method to identify impact craters, which detected 87% of the known impact craters. However, they only achieved $66.5 \pm 17\%$ of detection precision and 75% of F1 score. Silburt et al. [17] used the semantic segmentation algorithm U-Net [18] to segment and extract impact craters from a lunar elevation map with 56% accuracy of post-processing. Wang et al. [19] used ERU-Net to detect impact craters in a lunar elevation map, which improved 27.7% recall with respect to Silburt's method. However, the large number of parameters and slow recognition speed of ERU-Net disfavors real-time detection. It can be found that these previous deep learning works only used the digital elevation map. As known, there are mainly two kinds of lunar surface data: the digital elevation map (DEM) and the orthographic projection image derived from the Wide-Angle Camera (WAC) of the Lunar Reconnaissance Orbiter Camera. DEMs contain abundant morphological and topographical characteristics, and they are insensitive to illumination. However, DEMs have a weaker pixel intensity gradient between the rim and center, thus leading to intrinsic difficulty in identifying shallow craters [8]. Different from DEM, the visibility of impact craters in WAC is normally affected by the illumination angle. However, the WAC can keep the complex terrain context that is usually lost in DEM, but is "noisier" in this regard [20]. Due to different imaging conditions, some craters might be clearer in DEM than in WAC, and vice versa. Therefore, multiple image modalities can provide complementary information for better characterizing crater features than single-image data.

Motivated by the issue above, we, in this work, proposed feature complementary of DEM and WAC to more sufficiently capture the features of impact craters, in turn improving the detection performance. With the feature integration strategy, we accordingly explored a dual input convolutional neural network based on the U-NET structure, called Dual-Path. Dual-Path consists of three parts: a contracting path, bridging layers, and an expanding path in order to efficiently conduct feature complementary of the two images. In addition, an attentional mechanism and residual network are introduced to weight key information and avoid network degradation, respectively, and further improve the detection performance. The experimental result shows that Dual-Path can accurately identify the impact crater with a small number of parameters and has a faster inference ability, exhibiting higher accuracy than previous models on single elevation map data.

The rest of the paper is summarized as follows. Section 2 introduces our approach, including the data processing algorithm and object segmentation tasks in detail. In Section 3, we report the results and discuss the advantages of our approach. Section 4 summarizes this paper.

2. Methodology

In this section, we first introduce the processing method for the two types of images (DEM and WAC), and then describe the detailed architecture of Dual-Path. Finally, we apply the template matching algorithm to obtain the predicted impact crater and evaluate our results.

2.1. Data Preparation

In this work, we used two types of moon images. The digital elevation model (DEM) image was derived from the Lunar Reconnaissance Orbiter Camera (LROC) [21], while the orthographic projection image was obtained from the Wide-Angle Camera (WAC) of the Lunar Reconnaissance Orbiter Camera, which consists of eight sub regions [22]. Figures 1 and 2 show the two types of images. We set WAC region codes to be in the range of A–H regions (vide Figure 2) according to the longitude and latitude. The overall range of WAC data is consistent with DEM, both in the longitude range $[-60^\circ, 60^\circ]$ and latitude range $[-180^\circ, 180^\circ]$.

Figure 1. Digital elevation model in the longitude range $[-60^\circ, 60^\circ]$ and latitude range $[-180^\circ, 180^\circ]$ for the lunar image.

Figure 2. Eight regions of the WAC from the Lunar Reconnaissance Orbiter Camera in the longitude range $[-60^\circ, 60^\circ]$ and latitude range $[-180^\circ, 180^\circ]$ for the lunar image. Letters (A–H) denote different regions.

The DEM image used in the experiment had a resolution of 59 m/pix (512 pixels/degree), and the width and height of the entire DEM image were $184,320 \times 61,440$ pixels. The orthographic projection image resolution was 100 m/pix (303.23 pix/deg), and the width and height of each WAC image were $27,291 \times 18,194$ pixels. In order to ensure the consistency

of the DEM and WAC images, we used bicubic [23] down-sampling to obtain a new DEM image with a width and height of $92,160 \times 30,720$ pixels, and the adjusted picture resolution was 118 m/pixel (256 pixels/degree), which is same as that of the WAC.

As accepted, the shape of a crater will change with increasing crater diameter and age, and the resolution of the lunar image also affects the diameter of an impact crater. Thus, in order to predict impact craters with a wider diameter range, we adopted the strategy of random clipping, which specified the random cropped sub-picture size range as 500×500 pixels– 6500×6500 pixels. Then, the cropped sub-pictures were down sampled to 256×256 pixels, as the input picture size of Dual-Path was specified as 256×256 pixels. Its geographic range corresponded to 59×59 km– 767×767 km of the original geographic range, achieving a wide range. Similar to Siburt's work [17], we also focused on impact craters with a diameter of 10–80 pixels in the picture, through which we could extract the diameter of the impact crater in the range of 2304 m–239.684 km, achieving a wide range.

In addition, we used the existing two types of crater location information data to draw semantic segmentation annotations. The first data set came from the global crater dataset provided by Povilaitis et al. [24] using the original image data Lunar Orbiter Laser Altimeter (LOLA) and a digital terrain model (DTM) with a resolution of 64 pixels/degree, which included impact craters with diameters of 5–20 km. The second data set is the large-scale impact crater data set assembled by Head et al. [25], in which the diameters of the impact craters are greater than 20 km. The impact crater statistics of the two data sets are shown in Table 1.

Table 1. Lunar crater distribution of eight regions.

Regions	Longitude	Latitude	Head	Povilaitis	Number of Craters
A	(−180, −90)	(0, 60)	777	4403	5180
B	(−90, 0)	(0, 60)	210	904	1114
C	(0, 90)	(0, 60)	306	1234	1540
D	(90, 180)	(0, 60)	669	3544	4213
E	(−180, −90)	(−60, 0)	516	2579	3095
F	(−90, 0)	(−60, 0)	422	1559	1981
G	(0, 90)	(−60, 0)	666	2468	3134
H	(90, 180)	(−60, 0)	735	2644	3379
sum	(−180, 180)	(−60, 60)	4301	19,335	23,636

According to the longitude and latitude required, the Cartopy Python package [26] was used to convert the image into an orthographic projection format to minimize image distortion and make the edges of the impact crater in the image more rounded, as reflected by a comparison of Figure 3a–d. In addition, the image intensity was linearly adjusted to enhance the contrast. The label data were rings with a width of 1 pixel drawn based on the actual location of the impact crater. When marking an impact crater, its radius and center depended on the radius and center of the impact crater. Specifically, we circled all of the impact craters that existed in the Povilaitis and Head data sets with a 1-pixel-wide ring, as reflected by Figure 3e. Any impact crater with a radius of less than 1 pixel was not circled.

After data processing, we obtained 15,000 training data, 5000 verification data, and 5000 test data. Each datum contained a pair of DEM and WAC impact craters at the same position.

2.2. Dual-Path Network Structure

With the complementary strategy of the two types of images, we accordingly explored a novel U-Net-based convolutional neural network framework coupled with dual-path, called the Dual-Path network. In general, the U-Net structure is very good at processing images with a simple semantic structure, while the lunar image had simple semantics and a fixed structure [20]. In addition, many works [27–30] based on U-Net have already achieved great success in different fields, including the detection of impact craters [17]. Thus, in this work, we adopted the U-Net structure to explore the dual-path-based CNN

framework. It consisted of a dual contracting path, bridging layer, and expanding path. The dual contracting path individually extracted high-dimensional abstract features from the DEM and WAC data. The bridging layer was constructed to concatenate and transfer the information from the contracting path to the expanding path, in which a 1×1 convolution kernel was used to reduce the dimension of each layer in the contracting path. The expanding path recovered the feature map to the size of the original input image and restored the image information to obtain the segmentation result. Figure 4 shows the whole framework architecture of Dual-Path, and the sizes of the feature maps are shown in Table 2.

Table 2. Feature maps of the Dual-Path network structure.

Layer Name	Feature Maps (Input)	Feature Maps (Output)
Special Conv 1	$256 \times 256 \times 1$	$256 \times 256 \times 32$
Max Pooling 1	$256 \times 256 \times 32$	$128 \times 128 \times 32$
Special Conv 2	$128 \times 128 \times 32$	$128 \times 128 \times 64$
Max Pooling 2	$128 \times 128 \times 64$	$64 \times 64 \times 64$
Special Conv 3	$64 \times 64 \times 64$	$64 \times 64 \times 128$
Max Pooling 3	$64 \times 64 \times 128$	$32 \times 32 \times 128$
Special Conv 4	$32 \times 32 \times 128$	$32 \times 32 \times 256$
Max Pooling 4	$32 \times 32 \times 256$	$16 \times 16 \times 256$
Special Conv 5	$16 \times 16 \times 256$	$16 \times 16 \times 512$
Bridging with GC	$16 \times 16 \times 512$	$16 \times 16 \times 512$
Transpose Conv 5	$16 \times 16 \times 256$	$32 \times 32 \times 256$
Special Conv 6	$32 \times 32 \times 256$	$32 \times 32 \times 128$
Transpose Conv 6	$32 \times 32 \times 128$	$64 \times 64 \times 128$
Special Conv 7	$64 \times 64 \times 128$	$64 \times 64 \times 64$
Transpose Conv 7	$64 \times 64 \times 64$	$128 \times 128 \times 64$
Special Conv 8	$128 \times 128 \times 64$	$128 \times 128 \times 32$
Transpose Conv 8	$128 \times 128 \times 32$	$256 \times 256 \times 32$
Conv and Sigmoid	$256 \times 256 \times 32$	$256 \times 256 \times 1$

Figure 3. Some representative images from the preparation of DEM and WAC data. (a) Original WAC data after random cutting, (b) WAC after orthographic projection, (c) original DEM data after cutting, (d) DEM after orthographic projection, and (e) labels marked from the Povilaitis and Head datasets.

Figure 4. Architecture of the U-Net-based Dual-Path model proposed. It consists of a contracting path, bridging layer, and expanding path.

2.2.1. Contracting Path

The contracting path consisted of two independent input branches, which individually processed the DEM and WAC images, as illustrated by Figure 4. Each branch in the dual-path included five special deep convolution blocks (labeled as Conv Block), four max-pooling layers with a 2×2 pool size, and four dropout layers alternately. Figure 5a illustrates the detailed architecture of the special deep convolution block, which was composed of three 3×3 convolution cores and three BN + ReLU layers. The 3×3 convolution kernel used zero padding to ensure that the output size of the network model was consistent with that of the input image. BN + ReLU is batch normalization (BN) and the Rectified Linear Unit (ReLU).

Figure 5. Architecture of the special deep convolution block and Global Context Module in Dual-Path. $\oplus$ denotes broadcast element-wise addition and $\otimes$ denotes matrix multiplication.

As known, in the training process, we generally encounter inconsistent distributions, which would lead to slow training, gradient disappearance, and gradient explosion. Thus, it needs regularization to prevent overfitting. To this end, we utilized BN before the ReLU to avoid inconsistent data distribution and speed up the convergence. Furthermore, BN also reduced the influence of the front layer on the back layer so that the back layer could easier update the front layer [31]. In addition, to avoid network degradation and gradient disappearance, we introduced the residual module [32] into the special deep convolution blocks by skip-connecting the first 3×3 convolution with the third 3×3 convolution after the first layer BN + ReLU. Finally, the skip-connect result was sent to BN + ReLU to obtain the output of the special deep convolution block. Each special deep convolution block in layers 1, 2, 3, 4, and 5 contained 32, 64, 128, 256, and 512 filters, respectively, which were lighter than the previous U-NET [18] framework on impact craters by the number of filters in each convolution layer.

2.2.2. Bridging Layer and Expanding Path

The bridging layer was used to connect the contracting path and the expansive path such that transfers the information. In the bridging layer, we conducted a concatenation operation in every layer of the contracting path to fuse the feature graph from the two corresponding special deep convolution blocks in the dual input branches involving DEC and WAC, as illustrated by Figure 4. In the bridge path, 1×1 convolution was used to reduce the number of channels to half in order to accelerate the training speed and simultaneously weaken the aliasing effect of the upper sampling. As known, due to the down-sampling operation of max-pooling in the contracting path, the size of the feature map was reduced, resulting in less semantic information in the low-level feature map, but the target location was accurate. In contrast, the semantic information of the high-level feature map was stronger, but the target location was rough [33]. Thus, concatenation in every layer of the contracting path by the 1×1 convolution of the bridging layer could enhance the integrity of the location and semantics information, which was beneficial for

impact crater detection. The bridging layer from top to bottom contained 32, 64, 128, 256, and 512 filters.

The expansive path was used to restore images and obtain precise localization of impact rings, which was composed of one global context module, four transpose convolutions coupled with dropout, three special deep convolution blocks, and one Conv and Sigmoid layer, as shown in Figure 4. The special deep convolution blocks 6, 7, and 8 contained 128, 64, and 32 filters, respectively. Finally, the Conv and Sigmoid output layer outputted prediction results by a 1×1 convolutional layer with a sigmoid function. In the expansive path, we introduced the attention mechanism to further optimize the feature space. The attention mechanism of human vision is to scan the image quickly to obtain the target region that needs to be focused on. Similarly, the attention operation in deep learning can select the more critical information from a vast feature space to further improve the recognition performance. Thus, following Cao et al. [34], we constructed the global context in the expansive path, as shown in Figures 4 and 5b. Concretely, the features of all positions were firstly aggregated to form the context modeling. Then, the feature transform module was used to capture the correlation at the channel level. Finally, the global context features were merged into the features of all locations by addition. Layer normalization can simplify the optimization of two-tier architecture with bottleneck changes to obtain better performance. The global context module can improve the recognition speed and accuracy of the network without adding additional computation.

After the global context model, the optimized feature map was up sampled by transpose convolution, and then fused with the feature graph from the same layer in the bridging layer. Then, we further used the special deep convolution block to extract the features. Finally, the number of features was reduced to 1 by using the 1×1 convolution layer. After that, the sigmoid function was used to process the final output between 0 and 1. Based on this network structure, we obtained a pixel-level segmentation result with the same size as the original input image.

We used a dropout after each max-pooling and transpose convolution. Dropout can avoid network over fitting and accelerate network training by randomly removing some hidden neurons from the network during the training process [35].

2.2.3. Extraction of Impact Craters

The image predicted by Dual-Path was only the predicted value of pixels, so it was necessary to further extract the impact crater. Herein, we used the template matching algorithm in scikit-image [19,36] to extract the possible impact crater position from the predicted pixels. The matching threshold was set to 0.5 to obtain the circular matching position of the impact crater edge [17]. The extracted coordinates of the impact crater were recorded as (x_i, y_i, r_i) . The location of the impact crater marked by experts [24,25] was marked as $(\hat{x}_j, \hat{y}_j, \hat{r}_j)$. If the following matching formulae (Equations (1) and (2)) was satisfied, it was regarded as the impact crater; otherwise, it was regarded as the wrong impact crater.

$$\frac{((x_i - \hat{x}_j)^2 + (y_i - \hat{y}_j)^2)}{\min(r_i, \hat{r}_j)} < D_{x,y} \quad (1)$$

$$\frac{\text{abs}(r_i - \hat{r}_j)}{\min(r_i, \hat{r}_j)} < D_r \quad (2)$$

where $D_{x,y} = 1.8$, $D_r = 1.0$ are the values of the hyper-parameter in [17]. If we recorded the detected impact crater as 1 and the undetected impact crater as 0, then the impact crater detection task was transformed into a simple binary classification model.

2.3. Evaluation Metrics

For each impact crater, there are three possible results in the case of comparison between the predicted and the real results: true positive T_P , false positive T_n , and false negative F_n .

$$\text{Predict crater} = T_P + F_P + F_n \quad (3)$$

In order to evaluate the quality of the crater detection model, we used precision P and recall R to measure the accuracy of the model in terms of Equations (4) and (5).

$$P = \frac{T_P}{T_P + F_P} \quad (4)$$

$$R = \frac{T_P}{T_P + F_n} \quad (5)$$

Recall R and precision P are two contradictory measures. Generally speaking, when the recall rate is high, the precision rate is often low, vice versa. For example, if we hope to screen out as many impact craters as possible, we can achieve this by increasing the number of candidate impact craters. All the craters that are real impact craters will be selected. As a result, we can achieve high recall in this situation, but the precision will be low. If we want to select the real impact crater as much as possible, we only choose the most confident one. In this way, we can achieve high precision, but many real impact craters will be missed, resulting in low recall. Therefore, the F_1 scoring function is introduced to measure the balance between precision and recall, as shown in Equation (6).

$$F_1 = 2 \frac{P \times R}{P + R} \quad (6)$$

For application, the model with high recall can find more impact craters and produce more false-positive samples. However, for the obstacle avoidance requirements of celestial probes, the model with high recall can be selected in order to make the landing and the smooth operation on the planet's surface [37]. We introduce the general form F_β of F_1 measurement, as expressed by Equation (7).

$$F_\beta = \frac{(1 + \beta^2) \times P \times R}{(\beta^2 \times P) + R} \quad (7)$$

Here, β represents the relative importance of the recall rate over the precision rate. If $\beta = 1$, F_β is the standard scoring function F_1 . If $\beta < 1$, the weight of the precision rate is more influential. If $\beta > 1$, recall rate has a greater impact. Here, we chose $\beta = 2$ to more incline to a high recall rate (vide Equation (8)), hoping to find as many impact craters as possible.

$$F_2 = \frac{5 \times P \times R}{(4 \times P) + R} \quad (8)$$

For measuring the accuracy of the impact crater location, we referred to the measurement standard of Wang [19] and calculate the longitude error ($Error_L_o$), latitude error ($Error_L_a$), and radius error ($Error_R$) in terms of Equations (9)–(11).

$$Error_L_o = \frac{|L_o - l_o|}{(R + r)/2} \quad (9)$$

$$Error_L_a = \frac{|L_a - l_a|}{(R + r)/2} \quad (10)$$

$$Error_R = \frac{|R - r|}{(R + r)/2} \quad (11)$$

where L_o is the longitude value of the CNN-predicted crater and l_o is the longitude value of the corresponding ground-truth crater. L_a is the latitude value of the CNN-predicted crater and l_a is the latitude value of the corresponding ground-truth crater. R is the radius value of the CNN-predicted crater and r is the radius value of the corresponding ground-truth crater.

In order to measure the velocity of crater detection based on the neural network, we also proposed frames per second (FPS). FPS is a time-related concept, which represents the number of picture frames processed per second. The higher the FPS value, the more pictures the model detects per second, that is, the faster the detection speed, which is an index to evaluate the detection speed of the model.

3. Experiments and Results

In this section, we firstly verify the effectiveness of the complementary strategy of the two types of images and the impact of the global context module embedded in the model architecture. Then, we compare the performance of our Dual-Path model with some competitive models in order to evaluate the model's advantage. Finally, we test the robustness of our model on the whole moon. All experiments in this study were run on a server; the operating system was CentOS 7.5, Intel (R) Xeon (R) CPU E5-2630 v4@2.20GHz, the graphics card was an NVIDIA GeForce RTX 2080TI, 11 GB video memory, and the deep learning algorithm acceleration was carried out through CUDA10.0. All experiments were carried out under the Python 3.6.10 environment and Tensor-Flow 1.9.0.

3.1. Advantage of the Feature Complementary of the DEM and WAC Images

In order to validate the impact of the feature complementary of the DEM and WAC images on the model performance, we conducted a comparison between the dual image input and single image input. For the single image input, we deleted one branch of the Dual-Path network. Accordingly, the network was degenerated into the depth residual U-NET structure of a single path structure. For Dual-Path, we used the data described in Section 2.1, including 15,000 pairs of DEM and WAC pictures as the training set, and 5000 pairs of DEM and WAC pictures as the testing set. For the single path structure, we used the same data as Dual-Path, but only included DEM or WAC as input data. In the training process, the learning rate, batch size, loss function, and other parameters were the same for the three types of input data. Then, the best models were selected and applied to the testing set. Table 3 lists the comparison results.

Table 3. Experiments on single data input and dual data input derived from the individual best model.

Data Type ^a	Epoch Number ^b	Recall	Precision	F1-Score	F2-Score
DEM	30	74.3%	85.3%	78.1%	76.3%
WAC	6	68.9%	83.3%	73.5%	71.4%
DEM + WAC	24	85.0%	81.4%	82.1%	83.5%

^a DEM denotes only using DEM as input data; WAC represents only using WAC as input data; and DEM + WAC stands for the integration of DEM and WAC as input. ^b denotes the number of epochs needed to achieve the best model.

The number of epochs refers to the number of times that the learning algorithm will work through the entire training dataset. An epoch denotes one cycle through the full training dataset. It can be seen from Table 3 that the number of epochs needed to achieve the best model was different for the three types of input data due to their different complexities in data. The best model was obtained at the 30th epoch of training for DEM, while the WAC data achieved the best effect at the sixth epoch, as the DEM image was more complex than the WAC image. When combining the two types of images, the optimal model was reached at the 24th epoch, indicating that the introduction of WAC accelerated the network convergence.

Compared to the result from WAC as input data, the performance from DEM as input data was higher, as evidenced by the recall, precision, F1-score, and F2-score in Table 3. This result should be associated with illumination factors in the WAC data. As known, the orthographic projection image derived from the Wide-Angle Camera (WAC) is a perspective projection of cartography, through which the sphere is projected onto a secant plane or tangent plane. Consequently, the WAC image was dependent on the scanning time during imaging, involving the Wide-Angle Camera angle and the sun light angle, which would influence the appearance of the impact crater, such as the shadow region. In other words, the illumination factor in WAC would introduce complex shadow problems that are taken as ‘noisier’, in turn disfavoring the detection accuracy. It should be a main reason why previous studies on lunar segmentation networks almost all used the DEM image, rather than the WAC image. However, when we combined the two types of images to extract the features, the model performance greatly outperformed any single input, except for precision, as reflected by Table 3. Compared to DEM, the recall, F1-score, and F2-score were increased by 10.7%, 4%, and 7.2%, respectively, which were beneficial from the feature complementary. The result clearly demonstrates the advantage of image integration.

3.2. Ablation Experiment on Global Context Module

As outlined above, we introduced the global content (GC) module in the expansive path to further optimize the feature space, which utilized the attention mechanism. In order to evaluate the impact of the GC module on the model performance, we conducted an ablation experiment on the GC model by removing it from the Dual-Path. Table 4 shows the result of the ablation experiment. It can be seen that the performance of the Dual-Path was lowered upon removing the GC module. Despite the slight increase in the amount of network parameters after introducing the Global Context, the model performance was significantly improved. For example, the recall, F1-score, and F2-score were increased by 3%, 1.7%, and 2.4%, respectively. The result indicates that the Global Context Model could improve the recognition ability of the dual moon model in the case of not significantly increasing the number of parameters and calculations, confirming the rationality of our model construction.

Table 4. Ablation experiments on the Global Context (GC) module ^a.

GC Module ^a	Parameter	FPS	Epoch Number ^b	Recall	Precision	F1-Score	F2-Score
✓	12,418,562	36.099	22	85.0%	81.4%	82.1%	83.5%
×	12,351,617	35.857	21	82.0%	81.4%	80.4%	81.1%

^a ✓ denotes the Dual-Path model including GC module while × represents the Dual-Path without the GC module.

^b denotes the number of epochs needed to achieve the best model.

3.3. Comparisons with Other Competitive Methods

To further evaluate the detection performance of our Dual-Path model, we selected four competitive models to compare, including DeepMoon [17], ERU-Net [19], LinkNet [38], and U-Net [18]. DeepMoon and ERU-Net exhibited good performance in detecting the impact craters, which only used DEM as the dataset. The LinkNet [39–41] and U-Net [42–44] algorithms have been widely used in image segmentation. Thus, we took the four methods as competitive models. We set the number of starting filters to 112, which was generally used in the corresponding works. Following the related works, the four competitive models only used DEM as the dataset. Our Dual-Path model still adopted the dual image as the input (DEM and WAC). The same data split was used for all the models (15,000 training samples, 5000 verification samples, and 5000 testing samples). The comparison results are shown in Table 5.

Table 5. Model performance on the testing set for our Dual-Path model and four competitive models.

Metric	Deep Moon	ERU-Net	U-Net	LinkNet	Dual-Path
Parameters	10,278,017	23,740,305	40,989,313	38,105,425	12,418,562
FPS	34.9	14.1	7.7	8.4	36.1
Recall	43.3%	75.5%	63.7%	75.1%	85.0%
Precision	90.8%	86.8%	82.7%	87.8%	81.4%
F1-score	55.8%	79.5%	70.1%	79.6%	82.0%
F2-score	47.3%	76.8%	65.8%	76.6%	83.5%
Error _{L_o}	8.2%	7.3%	7.7%	7.3%	4.1%
Error _{L_a}	6.7%	6.9%	6.8%	6.8%	3.4%

Although DeepMoon was lowest in the amount of network parameters, its performance was the poorest, as evidenced by the recall, F1-score, and F2-score. Our Dual-Path model had slightly more network parameters than DeepMoon, yet our recall, F1-score, and F2-score were greatly increased by 41.7%, 26.2%, and 36.2%, respectively. The other three models presented much more network parameters than our Dual-Path model, about twice and triple. ERU-Net exhibited the best performance among the four competitive models. Compared with ERU-Net, our parameter amount was only half, but our recall, F1-score, and F2-score were increased by 9.5%, 2.5%, and 6.7%, respectively. In addition, our model had the highest FPS, indicating the fastest speed.

3.4. Robustness Testing on the Whole Moon

In order to further verify the robustness of our model, we used the Dual-Path network model to further detect other targets in eight different regions widespread over the whole moon (labelled as A–H in Figure 2), which were not included in our dataset above. According to the longitude and latitude, we randomly sampled 5000 images in each region and used the best Dual-Path model obtained to detect them. Table 6 lists the statistical results.

Table 6. Dual-Path test results in eight regions.

Regions	A	B	C	D	E	F	G	H
Craters in dataset	6215	5325	6798	6707	9564	1077	3262	5820
Restored craters	5103	4390	5501	5402	7011	917	2854	4857
Detected craters	6276	5172	6633	6456	8416	1136	3450	5957
New craters	1173	782	1132	1054	1405	219	596	1100
Omitted craters	1112	935	1297	1305	2553	160	408	963
Precision	81.3%	84.9%	82.9%	83.7%	83.3%	80.7%	82.7%	81.5%
Recall	82.1%	82.4%	80.9%	80.5%	73.3%	85.1%	87.5%	83.5%

As shown in Figures 1 and 2, the five regions labeled as A, D, E, G and H included more lunar land regions than the other regions, which were undulating and had high altitudes; thus, the image features were more complicated than the maria region and there were widespread impact craters. In contrast, the other three regions (labeled as B, C and F in Figures 1 and 2) included more lunar maria regions, as reflected by the darker color in Figure 2, where the impact craters were contiguously distributed without obvious shadow characteristics.

As can be seen from Table 6, the numbers of impact craters in the eight regions were quite different. However, the results detected by our model were very stable, in which the precision was in the range of 80.7–84.9% and recall was in the range of 80.5–87.5%, except for region E, with 73.3% of recall. The relatively low recall in region E should be attributed to the complex geological conditions and largely overlapping impact craters. However, its precision of 83.3% still ensured reliable detection of most impact craters, even in complex geological conditions. Region G achieved the highest recall (87.5%) due to the large proportion of lunar land region, simple terrain, and fewer impact craters, indicating that it was easier to recognize impact craters in the lunar land region. Although region

F contained the fewest craters (1077), the precision was lowest (80.7%). As reflected by Figure 2, region *F* included a large proportion of the lunar maria region and the impact craters were very sparse, which should have contributed to the relatively low precision and recall. For other regions, both the precision and recall were higher than 80%, further confirming the effectiveness of our model. Additionally, the result shows that the detection ability was slightly stronger in the lunar land region than that in the lunar maria region.

4. Discussion

As known, feature representation and model architecture are key factors to determining machine learning performance. Existing DL-based works on impact crater detection almost all used a single type of data, such as DEM or WAC. As outlined above, the two types of images characterized the impact crater from different, but complementary, aspects. Some craters might be clearer in DEM than WAC, and vice versa. Thus, the features derived from one single type of image data generally bring a risk of insufficient information, in turn disfavoring the detection performance. To alleviate this limitation, we proposed a feature complementary strategy by combining the DEM and WAC multi-source images for more sufficiently characterizing the impact features. In order to effectively conduct feature extraction and integration, we accordingly explored an advanced dual-path convolutional neural network (Dual-Path) based on a U-NET structure. As evidenced by some ablation experiments, the feature complementary significantly improved the detection performance with respect to the feature representation from the single image data. For the single image, it was not unexpected that the performance from DEM was superior to that from WAC. The comparison with four competitive models only using DEM featured further confirmed the advantage of feature complementary. In addition, our Dual-Path model presented the highest detection speed, as evidenced by the FPS in Table 5. These observations clearly show that the feature combination of DEM and WAC and the corresponding Dual-Path architecture can not only achieve high segmentation performance, but also a fast speed. The complementary strategy also provides guidelines for the application of deep learning in other fields. In addition, the independent testing on the whole moon showed satisfactory performance, almost higher than 80% for either the precision or recall metrics (Table 6), showcasing the robustness of our model to unseen cases and its good potential in practical application.

Despite the success that benefited from the feature complementary and model architecture, there were still some problems found from the detection result, as reflected by Figure 6 that representatively shows some detection results of our Dual-Path model. Specifically, Figure 6A(1)–E(1) shows the impact crater ground truth labels. Figure 6A(2)–E(2) is the segmentation results from the last Conv and Sigmoid layer of the Dual-Path model (vide Figure 4). Figure 6A(3)–E(3) and Figure 6A(4)–E(4) further show the final identification results after using template matching on DEM and WAC, respectively. As reflected by Figure 6, most of the impact craters were successfully recognized. However, some impact craters presented in Figure 6C–E were still missed, as highlighted in the red dashed boxes. For example, Figure 6C(2) shows a complex and dense crater scenario, leading to confusion in template matching. For the two impact craters closely connected in Figure 6D(2), they were merged into a large ring in the segmentation results, leading to a detection failure in Figure 6D(3–4). In addition, as shown in Figure 6E(2), the impact crater located on the edge of the image was easily expressed to be incomplete in the segmentation result, which contributed to its omission in the template matching. Thus, how to improve the template matching algorithm for the complex and incomplete segmentation deserves attention in the future, for example, using an adaptive threshold instead of a fixed threshold to extract craters as much as possible.

Figure 6. Some representative sample patches of the moon. (A(1)–E(1)) ground-truth labels from the raw images; (A(2)–E(2)) segmentation results from the Dual-Path; (A(3)–E(3)) detection results of the Dual-Path on DEM after template matching; (A(4)–E(4)) detection results of the Dual-Path on WAC after template matching. Blue circles represent correctly recognized craters that were successfully matched to the ground truth. Green circles denote new craters, while red circles represent unrecognized craters.

5. Conclusions

To address the feature limitation of a single type of moon image, we proposed the feature complementary strategy that combined DEM and WAC images. Accordingly, we explored a dual-path convolutional neural network based on the U-NET structure (Dual-Path model) to efficiently conduct feature complementary. The Dual-Path model consisted of a contracting path, bridging layers, and expanding path. The contracting path separately extracted features from the elevation map and orthographic projection images by means of two independent input branches, in which a special deep convolution block with a residual module was introduced to avoid network degradation and gradient disappearance.

The bridging layer integrated the elevation map and orthographic projection features by 1×1 convolution, which could reduce the number of parameters. Similar to the contracting path, the expanding path used the same special deep convolution block with a residual module to further fuse and learn the features from the bridge output and the feature map after transpose convolution. In addition, the attention mechanism was introduced to the expanding path to further optimize the feature space with the aid of a global context module. The experimental results demonstrated that the feature complementary strategy and the advanced dual-path architecture could effectively improve the detection performance with respect to any single image type. Our Dual-Path model trained on 15,000 elevation images and 15,000 orthographic projection images achieved 81.4% of precision, 85% of recall, and 83.5% of F2-score for the independent test set with the inclusion of 5000 images, superior to the four competitive models. In addition, our model was further verified by different regions on the whole moon, exhibiting high robustness and a fast speed, which is beneficial to application in the real-time monitoring of impact craters.

Author Contributions: Conceptualization, Y.M., R.Y. and Y.L.; methodology, Y.M. and R.Y.; software, Y.M., R.Y. and W.L.; validation, Y.M. and R.Y.; formal analysis, Y.M. and R.Y.; investigation, Y.M., R.Y. and W.L.; resources, Y.M. and R.Y.; data curation, Y.M., R.Y. and W.L.; writing—original draft preparation, Y.M. and R.Y.; writing—review and editing, Y.L.; supervision, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This project is supported by the Sichuan International Science and technology innovation cooperation project (Grant No. 2021YFH0140).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The author would like to thank the people who helped with the paper and the experiment.

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

References

1. NASA Catalogue of Lunar Nomenclature. Available online: <https://ntrs.nasa.gov/citations/19830003761> (accessed on 10 May 2021).
2. Losiak, A.; Wilhelms, D.E.; Byrne, C.J.; Thaisen, K.G.; Weider, S.Z.; Kohout, T.; Kring, D.A. A new lunar impact crater database. In Proceedings of the Lunar and Planetary Science Conference, Woodlands, TX, USA, 23–27 March 2009; p. 1532.
3. Urbach, E.R.; Stepinski, T.F. Automatic detection of sub-km craters in high resolution planetary images. *Planet. Space Sci.* **2009**, *57*, 880–887. [\[CrossRef\]](#)
4. Vijayan, S.; Vani, K.; Sanjeevi, S. Crater detection, classification and contextual information extraction in lunar images using a novel algorithm. *Icarus* **2013**, *226*, 798–815. [\[CrossRef\]](#)
5. Di, K.; Li, W.; Yue, Z.; Sun, Y.; Liu, Y. A machine learning approach to crater detection from topographic data. *Adv. Space Res.* **2014**, *54*, 2419–2429. [\[CrossRef\]](#)
6. Mu, Y.; Ding, W.; Tao, D.; Stepinski, T.F. Biologically inspired model for crater detection. In Proceedings of the 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 2487–2494.
7. Sawabe, Y.; Matsunaga, T.; Rokugawa, S. Automated detection and classification of lunar craters using multiple approaches. *Adv. Space Res.* **2006**, *37*, 21–27. [\[CrossRef\]](#)
8. Xie, Y.; Tang, G.; Yan, S.; Lin, H. Crater Detection Using the Morphological Characteristics of Chang'E-1 Digital Elevation Models. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 885–889. [\[CrossRef\]](#)
9. Chen, M.; Liu, D.; Qian, K.; Li, J.; Lei, M.; Zhou, Y. Lunar Crater Detection Based on Terrain Analysis and Mathematical Morphology Methods Using Digital Elevation Models. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3681–3692. [\[CrossRef\]](#)
10. Pedrosa, M.M.; Pina, P.; Machado, M.; Bandeira, L.; da Silva, E.A. Crater detection in multi-ring basins of mercury. *Lect. Notes Comput. Sci.* **2015**, *9117*, 522–529.
11. Wang, Y.; Wu, B. Active Machine Learning Approach for Crater Detection from Planetary Imagery and Digital Elevation Models. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5777–5789. [\[CrossRef\]](#)

12. Kang, Z.; Wang, X.; Hu, T.; Yang, J. Coarse-to-Fine Extraction of Small-Scale Lunar Impact Craters from the CCD Images of the Chang'E Lunar Orbiters. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 181–193. [\[CrossRef\]](#)
13. Stepinski, T.F.; Ding, W.; Vilalta, R. Detecting Impact Craters in Planetary Images Using Machine Learning. In *Intelligent Data Analysis for Real-Life Applications*; IGI Global: Hershey, PA, USA, 2012; pp. 146–159.
14. Jin, Y.; He, F.; Liu, S.; Tong, X. Small Scale Crater Detection based on Deep Learning with Multi-Temporal Samples of High-Resolution Images. In Proceedings of the 2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Shanghai, China, 5–7 August 2019; pp. 1–4.
15. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [\[CrossRef\]](#)
16. Ali-Dib, M.; Menou, K.; Jackson, A.P.; Zhu, C.; Hammond, N. Automated crater shape retrieval using weakly-supervised deep learning. *Icarus* **2020**, *345*, 113749. [\[CrossRef\]](#)
17. Silburt, A.; Ali-Dib, M.; Zhu, C.; Jackson, A.; Valencia, D.; Kissin, Y.; Tamayo, D.; Menou, K. Lunar crater identification via deep learning. *Icarus* **2019**, *317*, 27–38. [\[CrossRef\]](#)
18. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention, Cambridge, UK, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
19. Wang, S.; Fan, Z.; Li, Z.; Zhang, H.; Wei, C. An Effective Lunar Crater Recognition Algorithm Based on Convolutional Neural Network. *Remote Sens.* **2020**, *12*, 2694. [\[CrossRef\]](#)
20. DeLatte, D.; Crites, S.; Guttenberg, N.; Yairi, T. Automated crater detection algorithms from a machine learning perspective in the convolutional neural network era. *Adv. Space Res.* **2019**, *64*, 1615–1628. [\[CrossRef\]](#)
21. LRO LOLA and Kaguya Terrain Camera DEM Merge 60N60S 512ppd (59m). Available online: https://astrogeology.usgs.gov/search/map/Moon/LRO/LOLA/Lunar_LRO_LOLAKaguya_DEMmerge_60N60S_512ppd (accessed on 10 May 2021).
22. Lunar Reconnaissance Orbiter Camera Global Morphological Map of the Moon. Available online: http://wms.lroc.asu.edu/lroc/view_rdr/WAC_GLOBAL (accessed on 10 May 2021).
23. Keys, R.G. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* **1981**, *29*, 1153–1160. [\[CrossRef\]](#)
24. Povilaitis, R.; Robinson, M.; van der Bogert, C.; Hiesinger, H.; Meyer, H.; Ostrach, L. Crater density differences: Exploring regional resurfacing, secondary crater populations, and crater saturation equilibrium on the moon. *Planet. Space Sci.* **2018**, *162*, 41–51. [\[CrossRef\]](#)
25. Head, J.W.; Fassett, C.I.; Kadish, S.J.; Smith, D.E.; Zuber, M.T.; Neumann, G.A.; Mazarico, E. Global Distribution of Large Lunar Craters: Implications for Resurfacing and Impactor Populations. *Science* **2010**, *329*, 1504–1507. [\[CrossRef\]](#)
26. Cartopy: A Cartographic Python Library with a Matplotlib Inter-Face. Available online: <http://scitools.org.uk/cartopy/index.html> (accessed on 10 May 2021).
27. Iglovikov, V.; Shvets, A. Terausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. *ArXiv* **2018**, arXiv:1801.05746.
28. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 1st ed.; Springer: Granada, Spain, 2018; pp. 3–11.
29. Ding, P.L.K.; Li, Z.; Zhou, Y.; Li, B. Deep residual dense U-Net for resolution enhancement in accelerated MRI acquisition. In Proceedings of the Medical Imaging 2019: Image Processing, San Diego, CA, USA, 16–21 February 2019; SPIE: San Diego, CA, USA, 2019; Volume 10949, p. 109490F.
30. Qin, X.; Zhang, Z.; Huang, C. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognit.* **2020**, *106*, 107404. [\[CrossRef\]](#)
31. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 11 July 2015; pp. 448–456.
32. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 770–778.
33. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.
34. Cao, Y.; Xu, J.; Lin, S.; Wei, F.; Hu, H. GCNet: Non-Local Networks Meet Squeeze-Excitation Networks and Beyond. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Korea, 27–28 October 2019; pp. 1971–1980.
35. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
36. Boulogne, F.; Warner, J.D.; Neil, Y.E. Scikit-image: Image processing in Python. *PeerJ* **2014**, *2*, 453.
37. Delatte, D.M.; Crites, S.T.; Guttenberg, N.; Tasker, E.J.; Yairi, T. Segmentation Convolutional Neural Networks for Automatic Crater Detection on Mars. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 2944–2957. [\[CrossRef\]](#)

38. Zhou, L.; Zhang, C.; Wu, M. D-LinkNet: LinkNet With Pretrained Encoder and Dilated Convolution for High Resolution Satellite Imagery Road Extraction. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 182–186. [[CrossRef](#)]
39. Peng, B.; Li, Y.; Fan, K.; Yuan, L.; Tong, L.; He, L. New Network Based on D-Linknet and Densenet for High Resolution Satellite Imagery Road Extraction. In Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2019; pp. 3939–3942.
40. Zhu, Q.; Zheng, Y.; Jiang, Y.; Yang, J. Efficient Multi-Class Semantic Segmentation of High Resolution Aerial Imagery with Dilated LinkNet. In Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2019; pp. 1065–1068.
41. Yuan, S.; Yang, K.; Li, X.; Cai, H. Automatic Seamline Determination for Urban Image Mosaicking Based on Road Probability Map from the D-LinkNet Neural Network. *Sensors* **2020**, *20*, 1832. [[CrossRef](#)] [[PubMed](#)]
42. Alom, Z.; Yakopcic, C.; Hasan, M.; Taha, T.M.; Asari, V.K. Recurrent residual U-Net for medical image segmentation. *J. Med. Imaging* **2019**, *6*, 014006. [[CrossRef](#)] [[PubMed](#)]
43. Zhang, W.; Tang, P.; Zhao, L.; Huang, Q. A Comparative Study of U-Nets with Various Convolution Components for Building Extraction. In Proceedings of the 2019 Joint Urban Remote Sensing Event (JURSE), Vannes, France, 22–24 May 2019; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2019; pp. 1–4.
44. Ghosh, S.; Chaki, A.; Santosh, K. Improved U-Net architecture with VGG-16 for brain tumor segmentation. *Phys. Eng. Sci. Med.* **2021**, *44*, 703–712. [[CrossRef](#)]