

Article

A Multi-Scale Mapping Approach Based on a Deep Learning CNN Model for Reconstructing High-Resolution Urban DEMs

Ling Jiang ^{1,2,3}, Yang Hu ⁴, Xilin Xia ³, Qiuhua Liang ^{1,3,*} , Andrea Soltoggio ⁴ and Syed Rezwana Kabir ³

¹ State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China; jiangling_xs@163.com

² Anhui Engineering Laboratory of Geo-Information Smart Sensing and Services, Chuzhou University, Chuzhou 239000, China

³ School of Architecture, Building and Civil Engineering, Loughborough University, Leicestershire LE11 3TT, UK; x.xia2@lboro.ac.uk (X.X.); S.R.Kabir@lboro.ac.uk (S.R.K.)

⁴ School of Computer Science, Loughborough University, Leicestershire LE11 3TT, UK; yang.hu.1@warwick.ac.uk (Y.H.); a.soltoggio@lboro.ac.uk (A.S.)

* Correspondence: q.liang@lboro.ac.uk

Received: 11 April 2020; Accepted: 9 May 2020; Published: 12 May 2020



Abstract: The scarcity of high-resolution urban digital elevation model (DEM) datasets, particularly in certain developing countries, has posed a challenge for many water-related applications such as flood risk management. A solution to address this is to develop effective approaches to reconstruct high-resolution DEMs from their low-resolution equivalents that are more widely available. However, the current high-resolution DEM reconstruction approaches mainly focus on natural topography. Few attempts have been made for urban topography, which is typically an integration of complex artificial and natural features. This study proposed a novel multi-scale mapping approach based on convolutional neural network (CNN) to deal with the complex features of urban topography and to reconstruct high-resolution urban DEMs. The proposed multi-scale CNN model was firstly trained using urban DEMs that contained topographic features at different resolutions, and then used to reconstruct the urban DEM at a specified (high) resolution from a low-resolution equivalent. A two-level accuracy assessment approach was also designed to evaluate the performance of the proposed urban DEM reconstruction method, in terms of numerical accuracy and morphological accuracy. The proposed DEM reconstruction approach was applied to a 121 km² urbanized area in London, United Kingdom. Compared with other commonly used methods, the current CNN-based approach produced superior results, providing a cost-effective innovative method to acquire high-resolution DEMs in other data-scarce regions.

Keywords: urban DEM; high resolution; deep learning; convolutional neural network; multiple scales; flood modeling

1. Introduction

Digital elevation models (DEMs) have been widely used in many fields such as landform evolution, soil erosion modeling, and other geo-simulations [1–4]. In particular, DEMs provide indispensable data to support water resource management and flood risk assessment [5,6]. In urban flood risk assessment, the availability of high-resolution urban DEMs is crucial for the accurate representation of complex urban topographic features and is required for a reliable prediction of flood inundation to inform risk calculation [7,8].

The common ways of acquiring high-resolution urban DEMs include ground surveying and remote sensing through light detection and ranging (LiDAR) [9,10]. For LiDAR data in particular, many data filtering and fusion methods for improving data quality have been developed to support urban flood modelling to achieve better performance [11–15]. However, these LiDAR data processing methods are usually applied on high-resolution topographic datasets, and cannot create high-resolution DEMs from low-resolution data. Meanwhile, these data acquisition approaches are usually labor-intensive and financially expensive, hindering their wider application across a large domain. As such, high-resolution urban DEMs are not always available, especially for cities in developing countries. This essentially imposes a barrier for many applications including the development of effective urban flood risk management strategies that are necessary to be informed by high-resolution flood modelling results. Hence, it is necessary to develop alternative and more cost-effective approaches to construct high-resolution urban DEMs to support a wide range of applications.

Although high-resolution urban DEMs are not always available, low-resolution DEMs, on the other hand, are relatively easy to access. For example, there are a range of open-access global or regional DEMs, including Shuttle Radar Topography Mission (SRTM), ALOS World 3D, and pan-Arctic DEM [16]. Many relevant studies, such as CoastalDEM, show that these datasets provide important resources for water engineering applications including region-scale flood modelling and risk analysis [17,18]. However, the resolution of these open datasets is not sufficient to depict urban topographic features, including buildings and street networks, to support high-resolution flood modelling. Thus, it is desirable to develop effective techniques to enhance the quality of low-resolution DEMs to subsequently obtain high-resolution urban DEMs. Most of the existing high-resolution DEM reconstruction methods are developed for natural terrains, which may be generally classified into three categories: DEM interpolation, DEM enhancement, and learning-based DEM reconstruction.

The DEM interpolation methods, commonly including inverse distance weighting (IDW), bilinear interpolation (BI), cubic convolution (CC), and kriging interpolation (KI), are generally implemented according to spatial autocorrelation, that is, the correlation of the ground elevations between two points is inverse to the distance between them (also known as Tobler's first law of geography) [19–23]. These methods have been widely applied to generate high-resolution DEMs, but they commonly smoothen the fine topographic details (i.e., high frequency details) and lead to blurry information in the output products. To relax the limitation of these DEM interpolation methods, DEM enhancement methods are developed to restore the lost topographic features via introducing extra information to enhance the quality of low-resolution DEMs. The extra information may be derived from additional elevation points, contours, land-use maps, and flood extents [24–28], among others. DEM enhancement methods may improve the resolution and accuracy of DEMs by fusing multiple DEMs and datasets of different resolutions and from various sources. Nevertheless, the required extra-high-accuracy topographic information for the implementation of this type of method is still hard to acquire, especially for a large extent. The learning-based approaches generate high-resolution DEMs by establishing the correlation between low- and high-resolution DEMs through a training process [29–33]. Learning-based models can be trained to learn from multi-dimensional information, which may potentially produce high-resolution DEMs of better quality. However, less research has been done in this direction, and the existing learning-based models are relatively simple and not suitable for application in complex urban environments.

Most of the existing DEM reconstruction methods are developed and applied in natural terrains. Reconstruction of urban high-resolution DEMs faces extra challenges, and direct application of the existing methods in the complex urban environments is questionable and may not be feasible. Due to human interventions, urban topography is typically an intricate synthesis of natural and artificial features (e.g., roads, buildings, and different types of vegetation covers). For flood modelling, these key urban structures/features commonly define flood pathways and predominantly control the underlying hydrological and inundation processes, and must be accurately represented in urban DEMs to produce reliable simulation results [34–36]. Basically, the resolution of the topographic data must be consistent with the scale of the involved processes to ensure they can be reliably modelled and correctly interpreted [37]. Therefore, there is a strong research and practical need to develop new approaches to support multi-scale DEM reconstruction and efficiently reconstruct urban DEMs at a specified higher resolution from a low-resolution equivalent to support more accurate urban flood modeling and other applications.

Although cities are widely covered by artificial topographic features of different types and scales, they are planned and built according to specific regulations and codes. In other words, urban topography commonly presents a high level of self-similar structures or features, especially for cities in the same region. This is particularly suitable for the application of learning-based approaches. For example, convolutional neural network (CNN) [38,39] is a deep learning technique designed to automatically and adaptively learn the spatial hierarchies of image features and has been successfully applied in image recognition and many other fields, such as machine translation and autonomous driving [40–42]. An urban gridded DEM can be effectively regarded as an image. Using localized urban DEMs of different resolutions, a CNN model may be trained to recognize the patterns of topographic features varying from high to low resolutions or vice versa, and subsequently used to reconstruct high-resolution DEMs from the low-resolution data across a large area. Although it is challenging and expensive to create high-resolution DEMs across a large area covering an entire city, it is more feasible to acquire high-resolution DEMs in localized (small) areas using a range of survey techniques, such as an unmanned aerial vehicle (UAV). This paper presents an innovative multi-scale approach using a deep-learning CNN model to reconstruct high-resolution urban DEMs from a low-resolution dataset. To our best knowledge, this is the first attempt to construct a CNN-based multi-scale mapping framework for efficiently enhancing the resolution of urban DEMs, which may contribute to resolving the issue of data-scarcity for urban flood modelling and water engineering applications.

The rest of this paper is arranged as follows: Section 2 introduces the proposed multi-scale mapping approach for urban DEM reconstruction, followed by the introduction of a two-level accuracy assessment framework in Section 3; Section 4 describes the experiments undertaken to validate the proposed high-resolution urban DEM reconstruction approach; further discussion is given in Section 5; and finally several remarks are summarized in Section 6.

2. A CNN-Based Multi-Scale Mapping Approach

A multi-scale mapping approach based on CNN (MSM-CNN) was developed in this work to reconstruct high-resolution urban DEMs from a low-resolution dataset, which is illustrated in Figures 1 and 2. Herein, the low-resolution DEM is denoted as X , and the corresponding datasets at higher resolutions are denoted as Y^2, Y^4, \dots, Y^{2^n} , where the superscript 2^n indicates that these DEMs are at 2^n times higher resolution than the low-resolution DEM X , and n is a positive integer. The goal here was to reconstruct an urban DEM $F^{2^n}(X)$ at a higher resolution from the low-resolution DEM X to ensure that $F^{2^n}(X)$ was as close to the ground truth dataset Y^{2^n} as possible, which was achieved by training a CNN to learn mapping F .

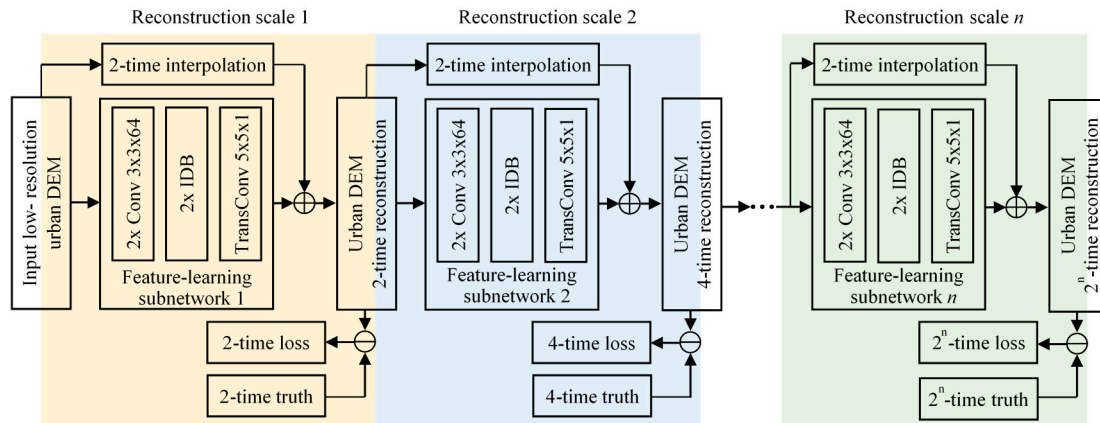


Figure 1. A multi-scale gradual network of multi-scale mapping approach based on convolutional neural network (MSM-CNN), in which IDB denotes the information distillation block as detailed in Figure 2; the symbols \oplus and \ominus represent the element-wise sum and loss-calculation operators, respectively; Conv and TransConv denote the convolutional and transposed convolution layers; the expression 2x Conv or 2x IDB represents two convolutional layers or two IDBs; and the size of convolutional or transposed convolution layers is in the format of width by height by number of filters (also referred as neurons or kernels), e.g., $3 \times 3 \times 64$.

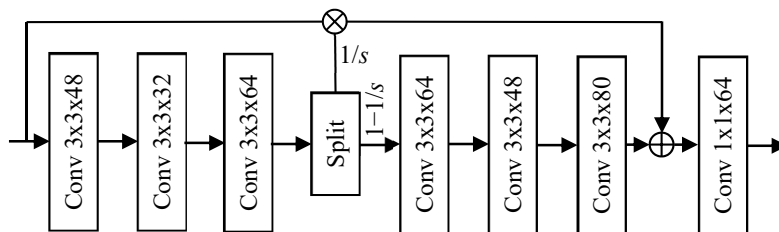


Figure 2. Architecture of information distillation block (IDB) in MSM-CNN, where s stands for the number of parts into which the feature maps are split, and \otimes is the concatenation operator.

2.1. Network Architecture

The detailed network architecture is shown in Figure 1, which consists of several subnetworks. Each of these subnetworks performs a 2-time reconstruction to its input urban DEM. According to the existing state-of-the-art results, a network with skip connections bypassing certain intermediate layers may lead to better performance [43–45]. Therefore, we introduced skip connections between the input and output of each of the subnetworks. Specifically, the input urban DEM of each subnetwork is interpolated to become two times its original resolution using a nearest neighbor (NN) method, and the interpolated data are then directly summed to the output of the feature-learning network. NN here was chosen due to its computational efficiency compared to other interpolation methods. The skip connections encourage the feature-learning networks to effectively learn and predict the missing topographic details from the low-resolution datasets to generate high-resolution datasets. Because each subnetwork only performs a 2-time reconstruction, the proposed architecture can effectively train a single network to construct urban DEMs at different higher resolutions.

In the proposed architecture, the feature-learning network is a key component in each of the subnetworks. Each feature-learning network starts with two convolutional layers with the kernel size specified as in Figure 1. The effect of the two convolutional layers is to extract initial features for further feature learning. The first two convolutional layers in the feature-learning network are followed by two information distillation blocks (IDBs) [46] to learn more powerful deep features for urban DEM reconstruction. The architecture of IDB is presented in Figure 2. The IDB starts with a stack of six convolutional layers, with the filter size specified as in Figure 2. After the first three layers in each IDB, the output feature maps are split into two parts. The $1-1/s$ percent of the feature channels are used as

the input to the next three layers, whereas the other $1/s$ percent feature channels is directly concatenated with the output of the next three layers. Such a structure creates skip connections and combines features in both shallower and deeper layers. The output of the first six blocks in IDB is passed to a seventh convolutional layer. This convolutional layer with 1×1 filters acts similarly to a bottleneck layer [47]; its effect is to combine and compress the shallow and deep features' output by the previous layers. Herein, although we used IDB as the backbone of the proposed network, other architectures could potentially also be used to replace IDB for feature learning. This paper focused on developing an innovative multi-scale network for urban DEM reconstruction rather than seeking the backbone architecture with the best performance; we selected IDB due to its reported excellent performance in accuracy and efficiency in computational cost. After the two IDBs, a transposed convolutional layer was applied to project the output feature maps of a subnetwork to a reconstruction at 2-time resolution with respect to the input of this subnetwork.

The proposed network uses rectified linear unit (ReLU) activation function, formulated as $y = \max(0, x)$, where x represents the input feature maps and y the output; y is equal to x if x is positive, otherwise y is 0. ReLU was adopted due to its widely reported effectiveness in the literature [43–45]. Herein, all of the convolutional layers are followed by a ReLU unless it is specifically mentioned otherwise.

A key advantage of the proposed multiple-scale architecture with respect to a single-scale architecture is that the multi-scale supervision was introduced to regularize the intermediate features of an urban DEM, which can faithfully enhance the output of each subnetwork to become as close to the high-resolution “true” DEM as possible. The adopted multi-scale supervision enables effortless and effective reconstruction of urban DEMs with enhanced accuracy at any specified higher resolution. Note that multi-scale design and computing losses at the intermediate network layers to guide the learning process have been widely used in deep neural network architectures [47–50]. In this paper, for the first time, we introduced this principle to the topic of urban DEM reconstruction.

2.2. Loss Function

The loss function used to train the network is based on mean absolute error (MAE). Let Y_i be the 2^i -time reconstruction result and R_i be the corresponding ground truth. The overall loss of network training denoted by MAE_{loss} is calculated as follows:

$$MAE_{\text{loss}} = \sum_{i=1}^n \left(\frac{1}{C} \sum_{j=1}^C |R_{i,j} - Y_{i,j}| \right) \quad (1)$$

where $R_{i,j}$ and $Y_{i,j}$ are the element in R_i and Y_i , respectively; C is the cell number; and n is the number of higher resolution datasets in the multi-scale gradual network.

Theoretically, a weighted sum could achieve better balance among the losses at different reconstructed resolutions. However, preliminary experiments reveal that the sum loss with equal weights is sufficient to achieve a good performance. We also compared the other metrics (e.g., mean squared error, structure similarity index, and peak signal-to-noise ratio) with MAE. MAE is not sensitive towards outliers and encourages less blurry surfaces, which is beneficial to reconstruct the spatial relationship between different artificial objects (e.g., roads and buildings) in urban terrains from the low-resolution data.

2.3. Network Training and Validation

We trained all the layers in the proposed network from scratch on the basis of the standard backpropagation with Adam optimizer [51] over a Caffe deep learning framework. The weights for convolutional layers were initialized using the method reported in [52]. The weight decay was set to 0.0001, and the learning rate was set to 0.0001 initially and reduced by a factor of 10 after 250 thousand iterations.

Prior to training the model, we prepared the training data by sampling it from the three selected training areas (see Section 4.1). Each of these sampled scenes had a spatial dimension of 500 by 500 cells and overlapped with neighboring scenes in both horizontal and vertical directions by 250 cells. The total number of sampled scenes available to train the model was 4107. A batch of 64 scenes was randomly selected from the sampled scenes that were from same training area, and then a patch from each scene was randomly cropped. These patches were then concatenated to form the batch of training data (i.e., we trained the model with a batch size of 64) during each forward–backward pass of the network. The size of a patch was chosen to meet the computational capacity, which depended on the number of scales in the network.

Upon successful completion of the training process, the first step was to examine whether the proposed method worked satisfactorily for scenes that had morphological characteristics similar to the training datasets. Therefore, a set of 456 scenes (not used during the training process) from the same three training areas was used to validate the model. However, investigating the generalization ability and transferability of the trained model in reconstructing high-resolution urban DEMs using spatially separated low-resolution data was more challenging. The effectiveness of the presented method over the test area is further analyzed in Section 4.

3. Two-Level Accuracy Assessment

To evaluate the performance of the proposed urban DEM reconstruction method, a two-level assessment approach was designed to quantify the numerical accuracy and morphological accuracy of the resulting products. Herein, the numerical accuracy is a quantification of elevation error at the cell locations, whereas the morphological accuracy is a region-scale quantification of morphology variance between the reconstructed urban DEM and ground truth.

3.1. Numerical Accuracy

Numerical accuracy was assessed by quantifying the difference of pointwise elevation between the reconstructed and “true” urban DEMs. Three metrics—MAE, root mean square error (RMSE), and standard deviation (STD)—were employed to quantify the numerical accuracy, which have been used as the standard statistical metrics for DEM vertical accuracy assessment [17,53]. The related equations to define RMSE and STD are given as follows:

$$\text{RMSE} = \sqrt{\frac{1}{c} \sum_{i=1}^c (x_i - y_i)^2} \quad (2)$$

$$\text{STD} = \sqrt{\frac{1}{c} \sum_{i=1}^c [(x_i - y_i) - (\bar{x} - \bar{y})]^2} \quad (3)$$

where c is the total count of valid grid cells, x denotes the ground elevations given by the reconstructed urban DEM, and y refers to the reference values.

3.2. Morphological Accuracy

A DEM not only represents the ground elevation at each of its cells, but also reveals the structure of the topography. As the skeleton of topography, topographic structure decides the spatial pattern of geomorphology [54]. Hence, the accuracy in representing the topographic structure is an essential indicator for DEM quality assessment. In the case of urban topography for application in flood or hydrological modelling, the topographic structure may be mainly reflected by the road networks and building clusters that have a significant impact on surface runoff and flow processes. Accordingly, the morphological accuracy, that is, the assessment of topographic feature difference, can be quantified by measuring the variances of the road profiles and building boundaries derived from the reconstructed urban DEM and the reference data.

The road-profile variance is measured through the following steps: (1) add vertices along each road centerline stepped by the cell size of the reconstructed urban DEM; (2) generate the road profiles respectively from the reconstructed and reference data; and (3) apply the Pearson's correlation coefficient (PCC) to quantify the variance between two profiles for each of the roads in the study area, and use the average and STD of PCCs to define the difference. Herein, the first two steps are implemented on the ArcGIS platform, and the last step is done using the Excel. The PCC is calculated as follows:

$$\text{PCC} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^m (y_i - \bar{y})^2}} \quad (4)$$

where m represents the number of the profile vertices, and x and y are the values corresponding to the reconstructed and reference profiles being compared.

On the ArcGIS platform, the variance of the building boundaries can be measured through three steps:

Step 1 is to consider the reference data by (1) preprocessing building polygons via merging the adjacent polygons and deleting those small and discrete patches according to an area threshold of 20 m², (2) obtaining the reference boundary line of each building patch and converting all lines to a raster aligned with the reference data, and (3) counting the boundary cells as the reference truth.

Step 2 is to extract building boundaries from the reconstructed urban DEM by (1) enhancing edge features (e.g., the boundary where a building meets a road) by a high-pass filter, (2) screening the candidates of boundary cells via an edge threshold of 1, and (3) obtaining the boundary cells using a thinning tool.

Step 3 is to quantify the variance by (1) selecting the boundary cells from step 2 according to the location of the reference boundary lines with no buffer, and buffers of 1, 2, and 3 times of the cell size of the reference data, respectively, and (2) calculating the ratio between the number of selected cells and that of the reference truth from step 1 successively. Finally, these four ratios are used to quantify the building-boundary variance.

4. Experiments and Results

In order to validate the performance of the proposed MSM-CNN method, a series of simulation experiments were undertaken. In the experiments, the MSM-CNN model was trained and applied to reconstruct high-resolution urban DEMs in the case study area. The experiments were performed on a single GPU (i.e., graphics processing unit) server with Nvidia K80 GPUs.

The produced outputs were compared with the results from several other popular interpolation or resample methods, including IDW, BI, CC, and KI. Herein, the ordinary KI was chosen due to its better accuracy among other types of KI for the study area. The experimental setup is illustrated in the flowchart shown in Figure 3. In the experiments, the urban DEMs at low resolutions of 2, 4, and 8 m were used to reconstruct high-resolution urban DEMs of 0.5 m to evaluate the performance of the multi-scale gradual network. It should be pointed out that, due to the lack of real datasets of 2, 4, and 8 m in the same period, we generated the three datasets by resampling 0.5 m data to ensure the consistency of evaluation benchmark. Herein, in the reconstructing phase, the test dataset was divided into scenes with a size of 250 by 250 cells (including an overlap of 125 cells with their neighbors), and finally, each scene was constructed individually and combined together to obtain the reconstructed urban DEM.

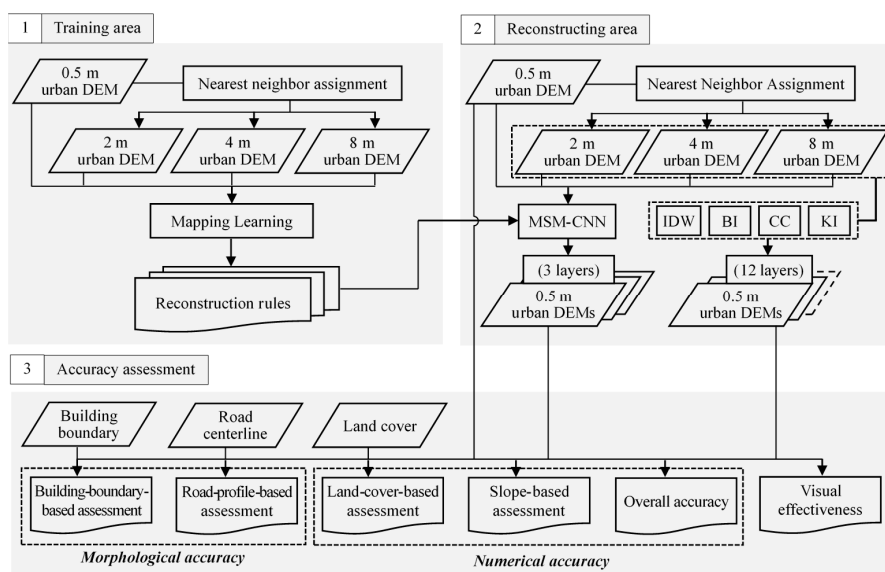


Figure 3. Experimental setup.

4.1. Study Area and Data

As one of the largest cities in the world, London, United Kingdom, is highly urbanized, with a population of 8 million, and was selected as the study area in this work. We firstly trained the MSM-CNN model using three small areas in the city. The three chosen training areas with significant different topographical features are located in the suburban, urban and rural regions, respectively. Each training site covered a 5 by 5 km area. After being trained, the MSM-CNN model was applied to reconstruct high-resolution DEMs in another larger area of 121 km², which is an urbanized area with mixed topographic features. The rationale to perform training and testing in different areas was that, although the overall urban designs could vary in different areas, the local features such as lines, edges, and blocks are similar across different natural and manmade structures; because a CNN focuses on local features, it could be used to reconstruct the urban structures in an area that is unseen in the training data. In this reconstructed area, eight samples of 1 by 1 km blocks were selected to facilitate morphological accuracy assessment. Figure 4 shows the locations of the training, reconstruction, and sample areas in the City of London.

In this work, a 0.5 m LiDAR DSM was used as the baseline high-resolution urban DEM, which is published by the Environment Agency, United Kingdom (<https://environment.data.gov.uk/ds/survey/index.jsp#/survey>). This dataset was employed for training the MSM-CNN model, and was used as the reference truth for assessing the reconstruction accuracy. The low-resolution DEMs for training and testing the MSM-CNN model were obtained from this 0.5 m DEM by resampling it to 2, 4, and 8 m resolutions using NN down-sampling (Figure 5). We selected NN instead of other alternative approaches such as BI or CC because this paper focused on urban DEM, which includes a large amount of abrupt elevation changes (e.g., a road with high buildings at both sides). For these specific types of data, methods such as BI and CC could be less suitable compared to NN, as they introduce “fake” elevation for the areas with abrupt features. Other relevant datasets of land cover, road centerline, and building were downloaded from Digimap (<https://digimap.edina.ac.uk>) for use in the current study. All of the above geospatial data were in the same coordinate reference system. It should be noted that if the coordinate reference systems of the essential data for MSM-CNN are different, geo-referencing (also known as image alignment) must be performed first.

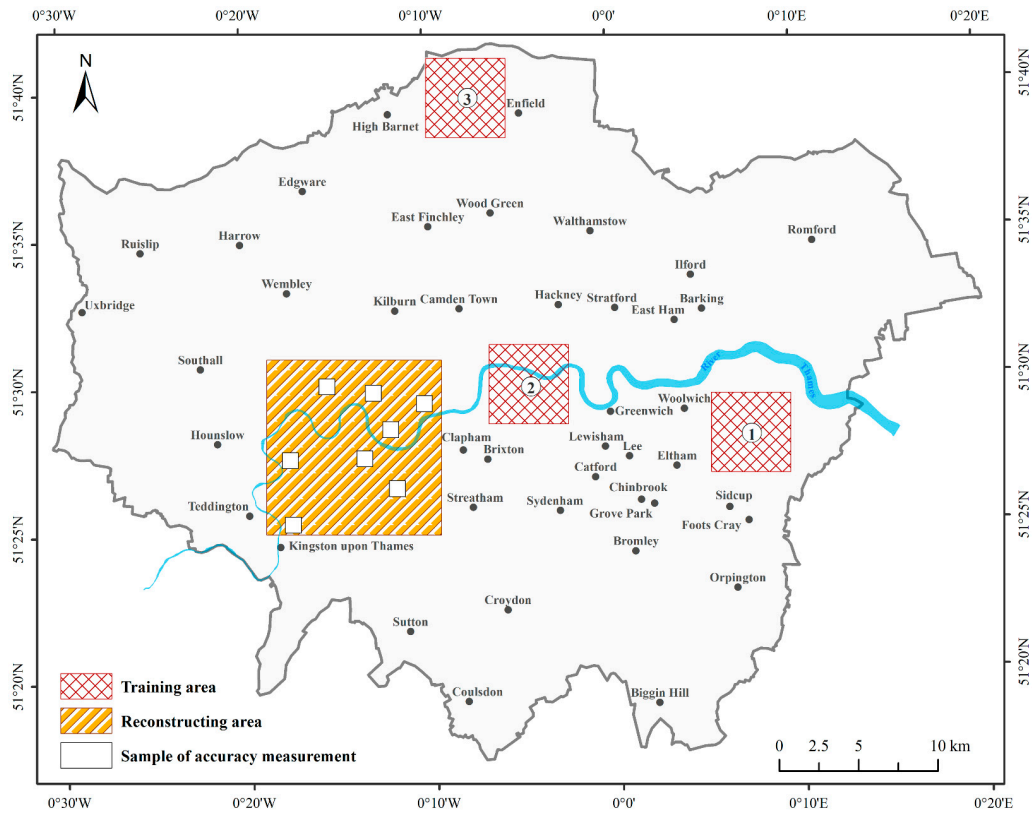


Figure 4. Location of study area.

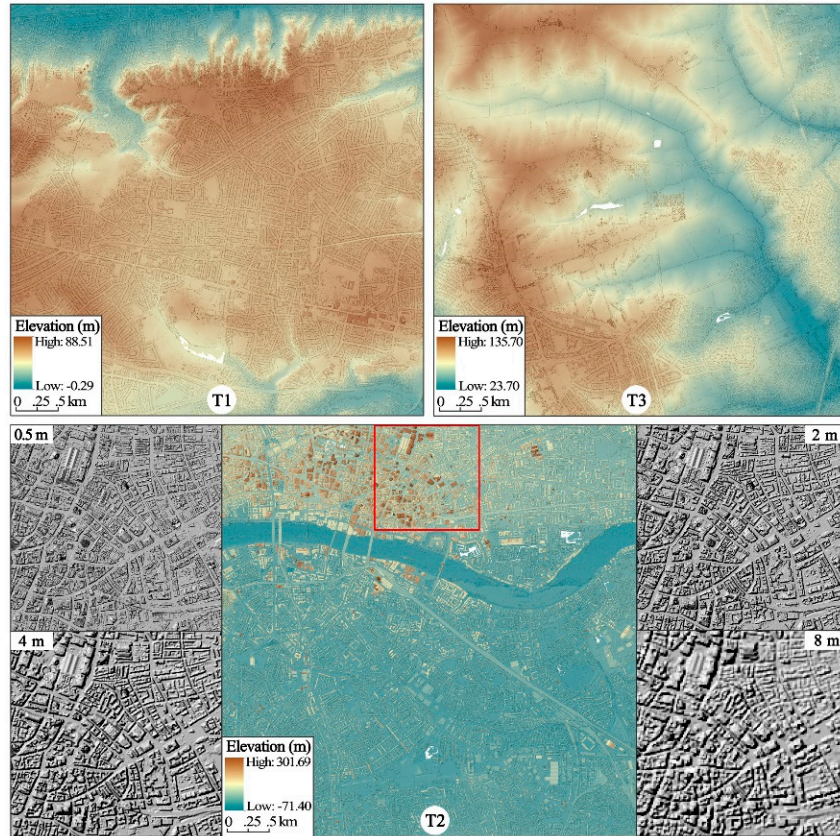


Figure 5. Urban digital elevation models (DEMs) of the training areas, where T1, T2, and T3 denote the training areas 1, 2, and 3, respectively.

4.2. Visual Assessment

The 0.5 m urban DEMs reconstructed using different methods were plotted together with the low-resolution counterparts of 8, 4, and 2 m in Figure 6. Naturally, the detailed features of urban topography were gradually lost as the resolution of the DEMs reduced from 0.5 to 2, 4, and 8 m (Figure 6a,e,i). The topographic structures related to road networks and building groups became blurry when the DEM resolution decreased. On the 8 m DEM, the roads and buildings became hard to identify. As depicted in Figure 6c,d,g,h,k,l, the BI, KI, CC, and IDW interpolation methods provided a certain level of enhancement in the topographic details. However, the level of enhancement was generally very limited, and in particular, it was not possible to restore most of the topographic structures from the lowest resolution (8 m) urban DEM. Moreover, hillock-like features were created in the three sets of the IDW reconstruction results, which did not conform to the morphological cognition of urban topography. It may be concluded that IDW is not applicable to urban topography, and IDW was therefore not chosen to support further accuracy assessment.

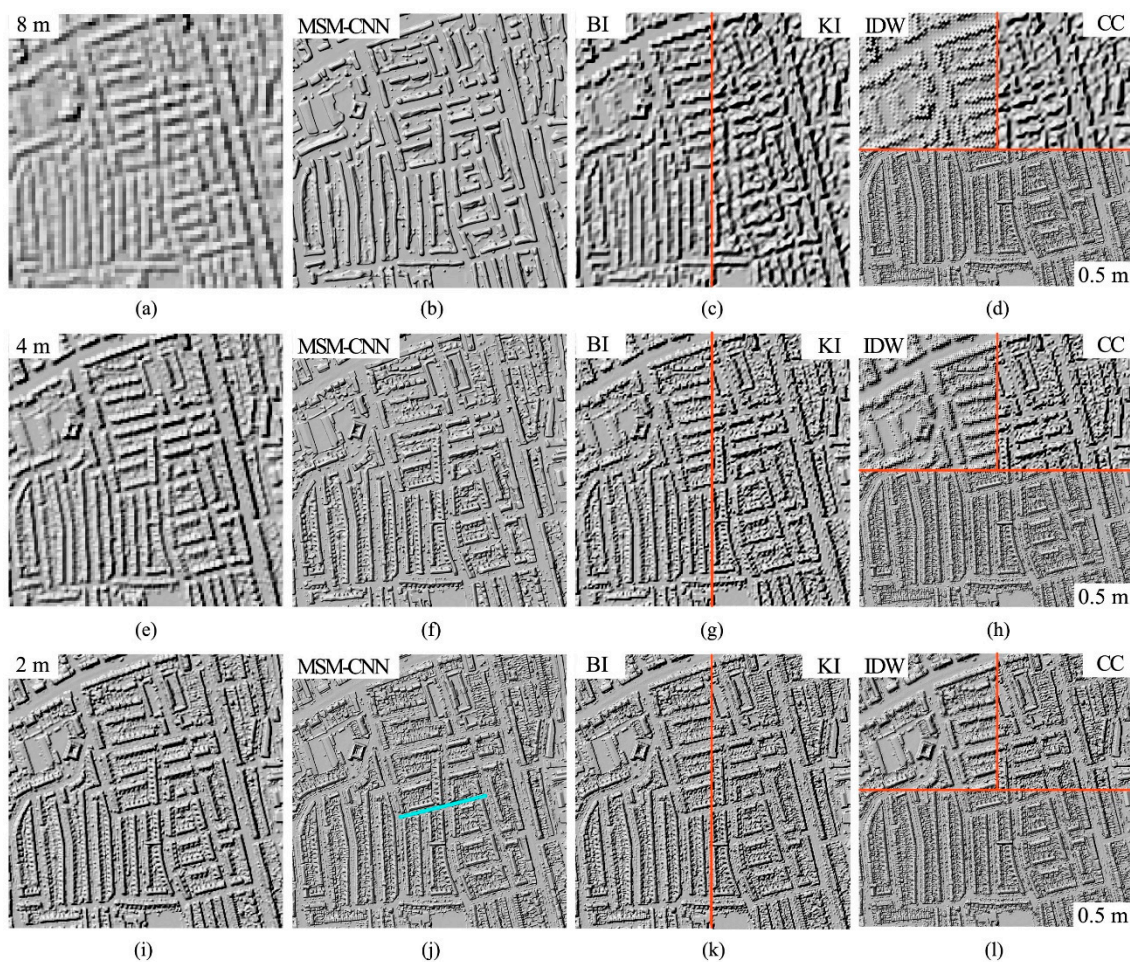


Figure 6. Reconstructed results in the study area (zoom-in): (a,e,i) the low-resolution urban DEMs of 8, 4, and 2 m; (b,f,j) the results reconstructed by MSM-CNN using the respective low-resolution DEMs at the same row; (c,g,k) from bilinear interpolation (BI; the left part) and kriging interpolation (KI; the right part); (d,h,l) results from inverse distance weighting (IDW; the upper-left part) and cubic convolution (CC; the upper-right part), and the reference urban DEM at 0.5 m (the bottom part). The highlight line in (j) is the road centerline for comparing reconstructed road profiles later.

The MSM-CNN evidently achieved better results for the reconstructions from all of the three low-resolution urban DEMs (Figure 6b,f,j). In the whole area, the topographic structure was restored

remarkably well, especially for the result reconstructed from the low-resolution DEM of 8 m, which showed good fidelity to the actual terrain. The MSM-CNN reconstructed DEM well represented both the continuous and abrupt features. Locally, the buildings and roads were clearly reconstructed, with their boundaries consistent with the reference terrain. As expected, the restored level of topographic details greatly depended on the input low-resolution urban DEMs, and more details were shown in the DEMs reconstructed from input datasets of higher resolutions. The results indicated that MSM-CNN can effectively achieve the multi-scale reconstruction to enhance the quality of low-resolution urban DEMs.

4.3. Numerical Accuracy

4.3.1. Overall Accuracy Analysis

Taking the original 0.5 m urban DEM as a reference, the results of numerical accuracy assessment of different reconstruction methods are listed in Table 1. From the 2 m low-resolution urban DEM, the 0.5 m product reconstructed by MSM-CNN was the most accurate, confirmed by the lowest MAE (0.194 m) and RMSE (0.918 m); meanwhile, the least accurate reconstruction result was obtained by CC, which had the highest MAE (0.234 m) and RMSE (1.028 m). The products reconstructed by BI and KI had the same MAE (0.234 m) but slightly different RMSEs of 1.012 and 1.019 m, respectively. From the lower-resolution DEM of 4 m, the best reconstruction result was still obtained by MSM-CNN, having MAE of 0.316 m and RMSE of 1.295 m. For the results reconstructed from the lowest-resolution dataset of 8 m, the MAE of the MSM-CNN reconstruction was slightly inferior to that of BI, but better than that of CC and KI; MSM-CNN also returned similar, but with slightly higher RMSE than BI and CC, and slightly lower value than KI.

Table 1. Accuracy statistics in the whole reconstructing area.

Low-Resolution Urban DEM	Method	MAE (m)	RMSE (m)	STD (m)
2 m	MSM-CNN	0.194	0.918	0.917
	BI	0.234	1.012	1.012
	CC	0.234	1.028	1.028
	KI	0.234	1.019	1.019
4 m	MSM-CNN	0.316	1.295	1.290
	BI	0.328	1.325	1.325
	CC	0.332	1.357	1.357
	KI	0.329	1.334	1.334
8 m	MSM-CNN	0.442	1.862	1.849
	BI	0.434	1.779	1.779
	CC	0.452	1.840	1.840
	KI	0.467	1.870	1.870

Overall, the numerical accuracy of the MSM-CNN reconstructions was mostly higher than that achieved by other interpolation methods. Meanwhile, it was noted that the variances of the numerical accuracy between MSM-CNN and other interpolation methods were not significant, which appeared to contrast with the visual comparison of the reconstruction results presented in Figure 5. The reason may be that the local elevation variation of urban topography in the reconstructing area was relatively small, and the overall statistics may not have efficiently reflected the small differences. It was therefore necessary to further investigate the performance of the MSM-CNN model by considering the morphological accuracy as well as conducting numerical accuracy assessment in groups, such as slope ranges and land covers.

4.3.2. Vertical Accuracy based on Slope Classification

We further investigated the vertical accuracy of the reconstruction methods by considering slope classification. The topographic features were divided into 10 ranges according to the ground surface

slopes, and then MAE and RMSE were respectively calculated for each of these ranges (Figure 7). Table 2 lists the average MAEs and RMSEs for all of the 10 slope ranges. Herein, the slope data were derived from the original 0.5 m urban DEM. From Figure 7a–c, a general increasing trend can be observed for both MAEs and RMSEs calculated for the different reconstruction results as the slope gradually increased. This indicated that the urban terrain relief as indicated by the slope factor had an obvious influence on the vertical accuracy of DEM reconstruction. As shown in Table 2, among all four approaches, MSM-CNN returned the highest accuracy confirmed by low RMSE and MAE for the reconstructions from all of the adopted low-resolution DEMs. The superior accuracy was maintained across all slope ranges until the slope was $\geq 100\%$, which covered 76% of the whole reconstruction area.

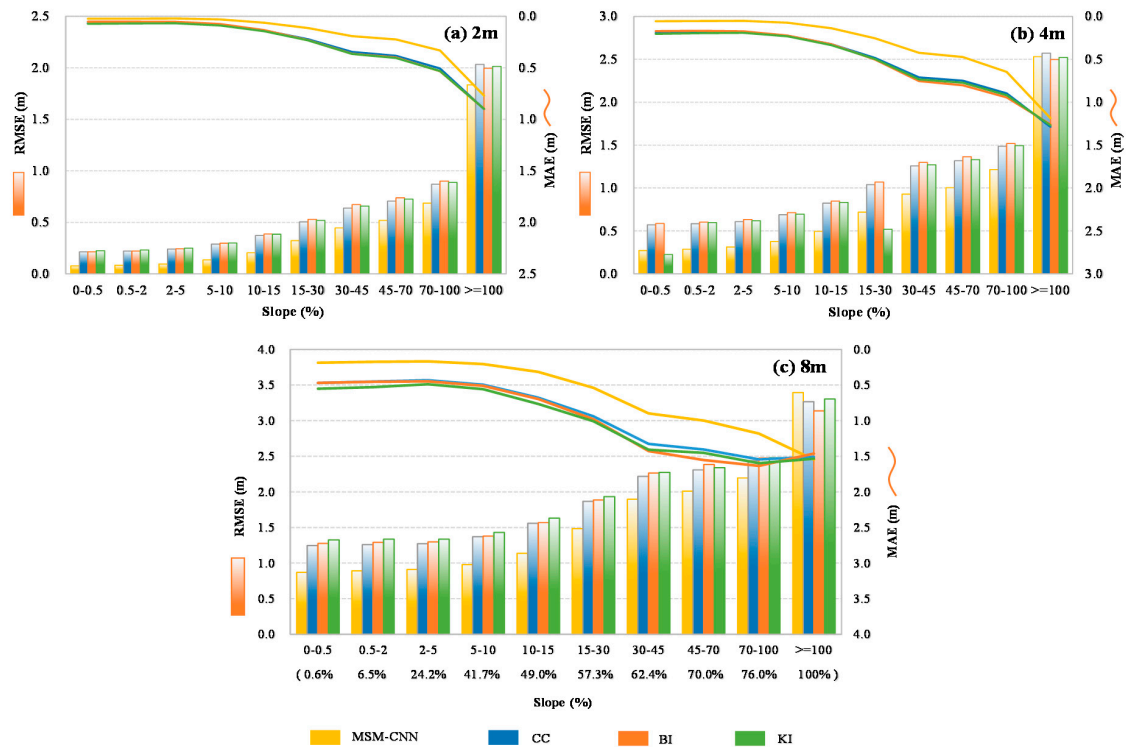


Figure 7. The accuracy statistics for different slope ranges in the whole reconstructing area: (a–c) root mean square errors (RMSEs) and mean absolute errors (MAEs) calculated for different DEMs reconstructed from the low-resolution 2, 4, and 8 m DEMs. The values inside the bracket below the x-axis in (c) are the accumulative frequency of each of the slope ranges.

Table 2. Average MAEs and RMSEs calculated for different slope ranges.

Low-Resolution Urban DEM	Method	Mean of MAE (m)	Mean of RMSE (m)
2 m	MSM-CNN	0.179	0.441
	BI	0.279	0.620
	CC	0.278	0.619
	KI	0.288	0.732
4 m	MSM-CNN	0.336	0.813
	BI	0.532	1.113
	CC	0.524	1.094
	KI	0.535	1.099
8 m	MSM-CNN	0.622	1.576
	BI	0.964	1.895
	CC	0.926	1.879
	KI	0.990	1.938

As the slope of the topography increased to $\geq 100\%$, both MAE and RMSE of the MSM-CNN reconstruction results were slightly higher than those of the other three methods when the reconstruction was conducted for the low-resolution DEM of 8 m. The MAEs of the BI, CC, and KI reconstruction results from the 8 m dataset started to decrease as the slope went beyond 100%, whereas their RMSEs continued to increase. In cities, the areas with the slope $\geq 100\%$ are mostly featured with abrupt change of terrain. Therefore, the reasons for the two aforementioned abnormalities may have been because the 8 m low-resolution urban DEM had smoothed out those sharp-fronted topographic features in this area, leading to the disappearance of the abrupt urban topography. As such, the MSM-CNN model may have exaggerated the reconstruction error by maximizing the restoration of the abrupt characteristic. For BI, CC, and KI, they essentially smoothed the abrupt terrain during the reconstruction without recreating abrupt change of the topography. Because the area featured with this highest slope range of $\geq 100\%$ took up 24% of the total area, the influence on the reconstruction results was evident. The findings may also explain the overall accuracy assessment result in Table 1, where the MSM-CNN reconstruction result from the 8 m DEM was slightly less accurate than those obtained using other interpolation methods.

4.3.3. Vertical Accuracy based on Land Cover Classification

For urban topography, terrain change is closely related to land cover types. Therefore, the vertical accuracy of the reconstructed DEMs from different approaches was also analyzed for various types of land covers. Herein, the urban land covers were divided into five types for analysis, including roads (RD), buildings (BG), natural environments (NT), multi-surfaces (ME), and others (OR). NT included those areas representing geographic extents of natural environments and terrains. ME comprised all of the artificial surfaces that are mainly around buildings, such as yards and plazas. Except for the first four types, the rest were classified as OR. Figure 8 illustrates the distribution of different land covers in a sample area within the case study site.



Figure 8. Different land covers in a sample area inside the reconstruction site.

Figure 9 shows the statistics of MAE and RMSE across different land covers for each of the reconstructed DEMs. For all of the land cover types, MSM-CNN returns smaller MAEs than all other alternative approaches for all of the reconstruction experiments. However, for NT, the MSM-CNN products reconstructed from the 4 and 8 DEMs only gave slightly higher RMSE than the results produced by BI. This again demonstrated that MSM-CNN is well applicable to both natural and artificial terrain in urbanized cities, whereas the interpolation methods were more suitable for application to natural terrain, and did not produce favorable results for urban topography. It is interesting to note that for land cover types of RD and BG, the MAEs of the MSM-CNN DEMs reconstructed from all three low-resolution DEMs were much smaller than other reconstruction results. Obviously, these were the two major land cover types in the urbanized areas and covered approximately 40% of the total area in the current study site. The performance analysis results effectively demonstrated that the current MSM-CNN approach offered better capability in restoring urban topographic structures with a high fidelity. In addition, the errors calculated for ME were relatively high for all reconstruction results, although the corresponding topography inherently had a low relief. A possible reason may have been that vegetation was not removed from the original 0.5 m urban DEM created from LiDAR data. Vegetation cover may have significantly affected the reconstruction accuracy because its elevation changed disorderly and behaved like random noise, which is difficult to be reliably reconstructed from low-resolution DEMs.

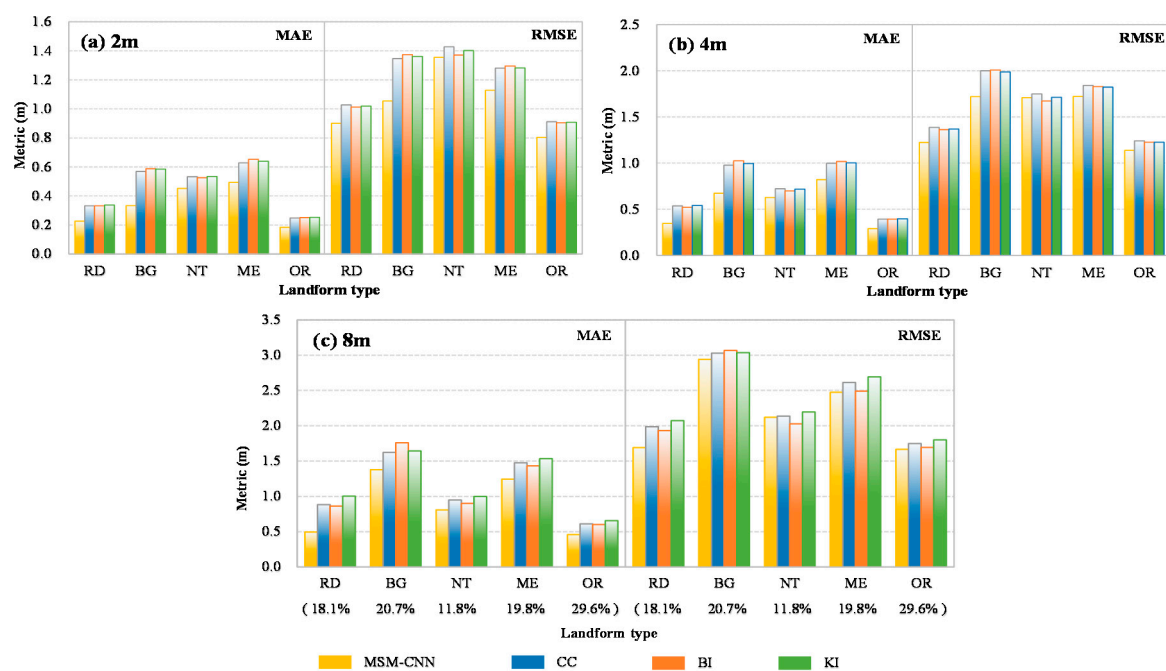


Figure 9. The accuracy metrics calculated for different land covers: (a,b,c) MAEs and RMSEs for different DEMs reconstructed from the 2, 4, and 8 m DEMs. The bracketed numbers below the x-axis in (c) indicate the frequency of each land cover.

4.4. Morphological Accuracy

4.4.1. Accuracy Assessment Based on Road Profiles

Figure 10 illustrates the centerline profiles of a road extracted from different reconstructed DEMs. The location of the selected road section is shown in Figure 6j. Obviously, the detailed features of urban topography were gradually lost as the resolution of DEMs reduced from 0.5 to 2, 4, and 8 m (Figure 6a,e,i), leading to blurry topographic structures related to road networks and buildings. Comparing the results obtained using different reconstruction methods, the MSM-CNN road profiles reconstructed from all three lower-resolution urban DEMs showed great agreement with the reference

profiles extracted from the original 0.5 m dataset. On the contrary, the road profiles generated by BI, CC, and KI showed spurious oscillations that were inconsistent with the morphology of urban roads. In particular, for the reconstructed results from the lower-resolution 4 or 8 m urban DEMs, the oscillations in the BI, CC, and KI products were so strong that the centerline profiles were no longer recognizable as a road. The potential reason for these results may have been that the BI, CC, and KI interpolation methods were implemented according to the spatial correlations between neighbors, whereas MSM-CNN was performed by the learned multi-dimensional patterns of topographic features varying from the high to low resolutions. When the DEM resolution decreased, the cell location where the prediction was being made had weaker or no clear spatial correlation with its neighbors. As such, the three CC or KI road profiles unexpectedly showed many deep ditches, which were again inconsistent with normal urban road morphology. The results confirmed the superior capability of the proposed MSM-CNN model in reliably reproducing urban morphology.

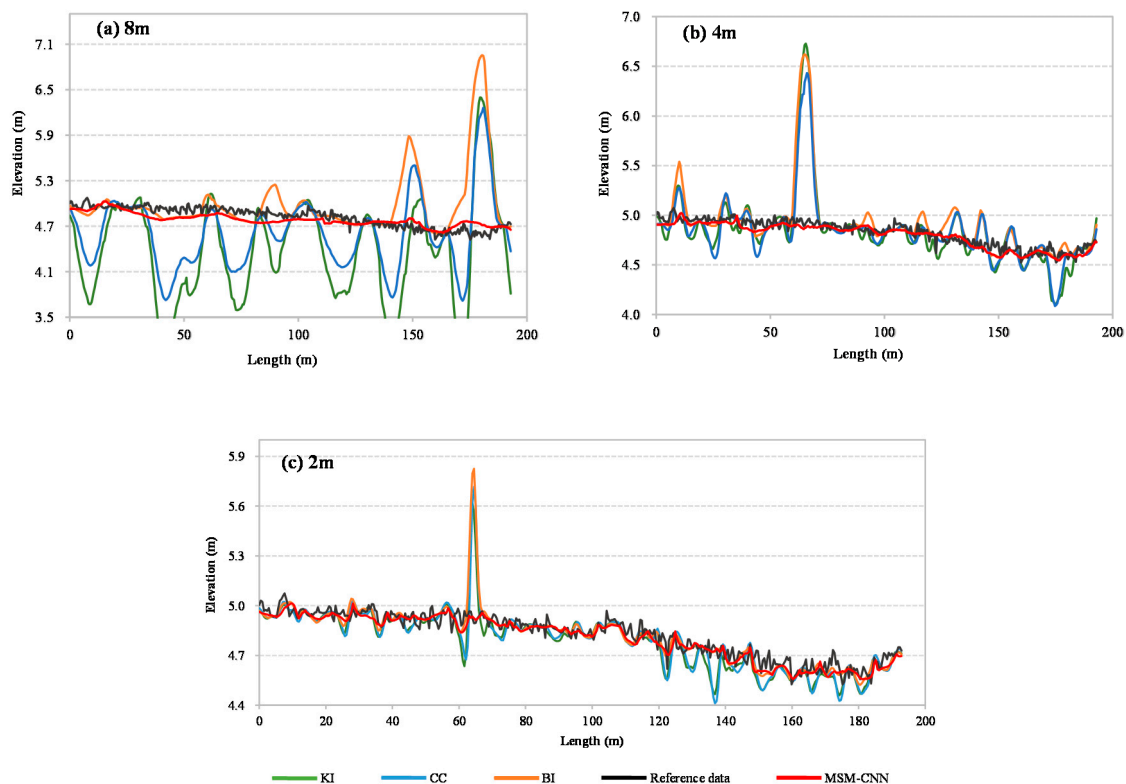


Figure 10. Road profiles extracted from the DEMs reconstructed using different methods: (a–c) the road profiles from the 8, 4, and 2 m reconstructed DEMs, respectively.

On the basis of the previous accuracy assessment results, BI produced better reconstruction results than the other two interpolation methods. Therefore, the following analysis was focused on comparing the morphological accuracy between the MSM-CNN and BI reconstruction results. Table 3 summarizes the statistics of the road-profile variance to quantify the morphological accuracy of the results. For the 4-time reconstructions (i.e., the 0.5 urban DEMs reconstructed from the 2 m equivalent), MSM-CNN clearly gave a better result than BI. According to the PCCs calculated for the reconstructed road profiles, 51% of the MSM-CNN reconstructed profiles had a PCC greater than 0.95, whereas only 38% of the BI reconstructed profiles reached the same level. For the MSM-CNN and BI reconstructions from the 4 m urban DEM, the difference in the morphological accuracy was significantly increased, as indicated by the average PCC of 0.79 for the MSM-CNN profiles and 0.66 for the BI profiles. Although 51% of the MSM-CNN reconstructed road profiles had the PCC greater than 0.9, only 29% of the BI profiles were able to reach this level. For 16-time reconstruction, that is, reconstructing the urban DEMs from 8 m coarse resolution to 0.5 m fine resolution, the improved morphological accuracy achieved by

MSM-CNN became even more prominent, and an improvement of 42% was achieved when compared with BI. The results demonstrated that the advantage of MSM-CNN in improving the morphological accuracy as represented by road-profile variance became more distinct as the resolution of the input urban DEM became coarser. In summary, the MSM-CNN reconstruction could substantially enhance the quality of low-resolution urban DEMs through improving morphological accuracy.

Table 3. Morphological accuracy statistics of road-profile variance.

Low-Resolution Urban DEM	Method	Mean of PCC	STD of PCC
2 m	MSM-CNN	0.89	0.15
	BI	0.83	0.20
4 m	MSM-CNN	0.79	0.24
	BI	0.66	0.30
8 m	MSM-CNN	0.68	0.33
	BI	0.48	0.36

4.4.2. Accuracy Evaluation Based on Building Boundary Reconstruction

Using the extraction method described in Section 3.2, building boundaries were delineated from the MSM-CNN and BI reconstructed DEMs for comparison, as shown in Figure 11, in which the reference boundary data are also presented in the vector format. As shown in Figure 11a for the 16-time reconstructions, the overall shapes of the boundaries were reasonably well reproduced by MSM-CNN, although certain fine-level details were smoothed out, which was as expected. However, almost no building boundary could be detected from the BI reconstructions. Figure 11b illustrates the reconstructions from the 4 m DEM. MSM-CNN representation of building boundaries was further improved and building corners could be clearly recognized. However, BI still failed to reconstruct the overall shape of the building boundaries. As exhibited in Figure 11c, the building boundaries in the MSM-CNN product reconstructed from the 2 m urban DEM were continuous and close to the reference, whereas the building boundaries produced by BI were typically segmented and did not align well with the reference. Evidently, MSM-CNN outperformed BI in restoring detailed features of urban topography and was more suitable for urban applications.

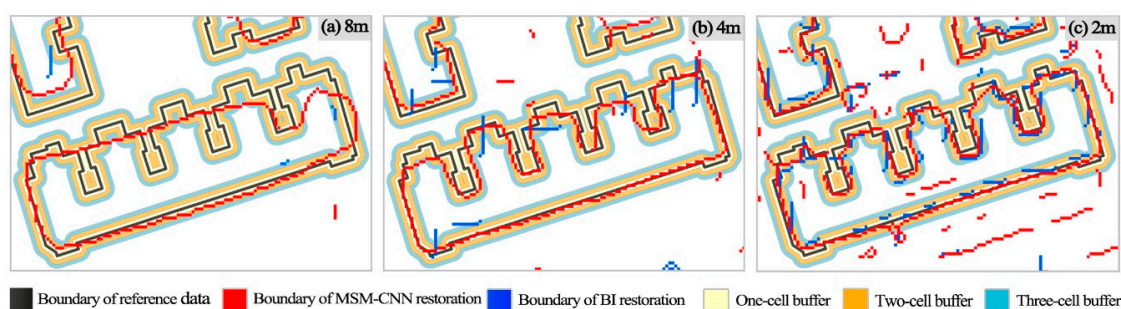


Figure 11. Building boundaries extracted from the reconstructed DEMs: (a,b,c) DEMs reconstructed from the 8, 4, and 2 m DEMs, respectively. One-cell, two-cell, and three-cell buffers are respectively the zones with widths of 1, 2, and 3 times the reconstructed cell size around the reference boundary lines.

To quantify the morphological accuracy of building boundary reconstruction, the percentage of correctly restored boundary cells was calculated and plotted in Figure 12. Overall, compared with BI, MSM-CNN presented clear superiority, especially for the reconstructions from lower-resolution DEMs. As expected, regardless the method being used, the morphological accuracy was calculated to be the highest for the 4-time reconstructions for each of the buffer ranges, followed by 8-time and 16-time reconstructions. The accuracy evaluated for the 4-time and 16-time MSM-CNN reconstructions only differed by an average of 2.5 times for the four buffer ranges. However, the accuracy difference

unexpectedly reached 16.2 times for the corresponding BI reconstructions. When the buffer distance was chosen as three cells (approximately 2 m where the cell size was 0.5 m), the percentage of correctly restored boundary cells returned by MSM-CNN was 70.23% for the 4-time reconstruction, and 34.52% for 16-time reconstruction where the resolution of the input DEM (8 m) was nearly four times larger than the buffer distance. For BI, only 42.73% of the boundary cells were correctly restored by the 4-time reconstruction; for 16-time, the figure substantially dropped to only 2.91%. This effectively demonstrates that MSM-CNN consistently outperformed BI in restoring building details.

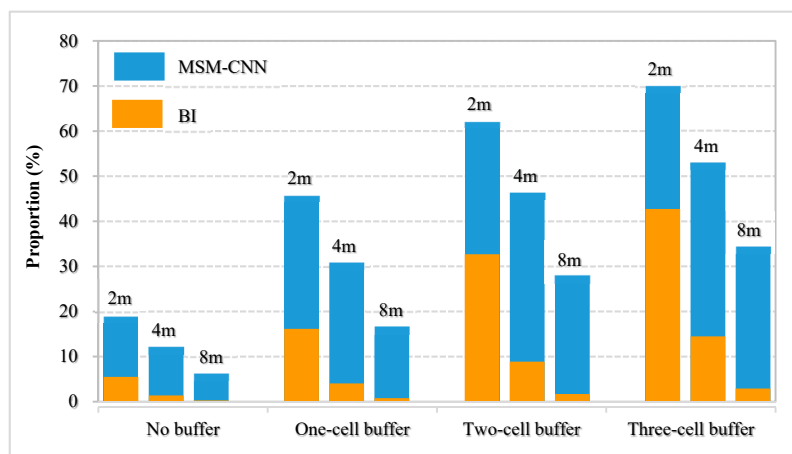


Figure 12. Morphological accuracy statistics of the building boundary reconstructions. The label of 2, 4, or 8 m on the top of each column denotes that the DEMs are reconstructed from the 2, 4, or 8 m low-resolution DEMs.

5. Discussion

5.1. Factors Influencing the Performance of DEM Reconstruction

5.1.1. Training Data

In the current selected case study area, the spatial topographic patterns of the reconstructed area resembled those of the test blocks. However, this might not be the case if the spatial heterogeneity of the test area is increased, and thus the training areas and data should be chosen with care. It has been widely recognized that the quality of training data has a major influence on the performance of a deep learning model [38,55]. For MSM-CNN, the reconstruction accuracy is potentially influenced by three factors: (1) typicality, (2) coverage, and (3) scale of the training data. Typicality requires that the training data should represent the typical features of urban topography to be reconstructed. Ideally, the training data should cover typical sample areas of the reconstructing site. In other words, it might be possible to increase the generalization capacity of the model by enriching the training database through diversifying the samples (e.g., adding more scenes from different cities). To rapidly acquire high-accuracy topographic data in these small and typical areas, the UAV photogrammetry is now entirely competent [56–58]. In theory, the larger area the training dataset covers, the more features of the urban topography can be learned. Nevertheless, the use of larger coverage of training data inevitably increases the cost in obtaining the sample datasets to train the learning model. Therefore, it is necessary to find a balance between the reconstruction accuracy and the coverage of training datasets. Alternatively, the model can be improved by leveraging a transfer learning method. In this approach, one can retrain a pre-trained model and only use new samples for which the model has not been trained before. This eliminates the need for developing and training a new model from scratch for regions with different topographical features.

For the implementation of the proposed multi-scale reconstruction approach, this work applied NN down-sampling to produce the low-resolution urban DEMs. Although the NN-based down-sampled

data can validate the current MSM-CNN, the effect of different down-sampling methods should be further investigated, and it would be better to collect and use real low-resolution datasets if available. Meanwhile, the range between the lower and upper resolutions for training was also better to cover the target range for high-resolution reconstruction.

5.1.2. Enhancement with Additional Terrain Information

On the basis of the quantitative assessment approaches designed and used in this work, it was indicated that the reconstruction accuracy varied with the land covers, slope ranges, and details of artificial buildings. This implies that the features of urban DEMs may be better learned by including additional terrain information to improve reconstruction quality. For example, land covers provide dominant features of urban topography. Land cover types may be considered in the learning process by distinguishing different types of topographic features, such as buildings, roads, water surfaces, and natural environments (i.e., natural terrains with relatively high relief). With the advanced image classification techniques, the high-resolution remote-sensing imagery is fully capable of mapping the above-mentioned land covers. Terrain attributes, such as slope, curvature, or roughness, define the multi-dimensional features of urban topography and may be also considered to improve the proposed deep learning process. These attributes can be straightforwardly derived from the corresponding urban DEMs; once the multi-layer attributes are classified, the weight of each layer may also be considered to facilitate a better learning process. Semantic knowledge is another source of information that may be considered. Herein, topographic semanteme refers to the rules of urban constructions, for example, the transversal and longitudinal gradients of roads. The semantic knowledge may be utilized to refine the urban topography. Overall, the present MSM-CNN model can be further improved to accommodate more topographic information to further enhance its performance, which deserves attention in future research.

5.2. Accuracy Assessment of Urban DEMs

In regard to urban DEMs, vertical accuracy assessment is a critical step to ensure their quality and support their further applications. The experimental results effectively demonstrate that accuracy assessment of urban DEMs must consider both the numerical and morphological accuracy. Herein, we propose the concept of morphological accuracy and present two basic indicators. It is worthy to strengthen the morphological indicators according to the morphological characteristics of different terrain objects and their spatial relationships. Moreover, approaches to combine the indicators of the numerical and morphological accuracy for integrated assessment should be further explored, such as using weighted sum. When developing integrated assessment methods, it is worth considering that these indicators probably have different dimensions and orders of magnitude.

5.3. Application of MSM-CNN in Water Science and Engineering Fields

The aim of developing the proposed MSM-CNN is to provide a feasible approach to reconstruct high-resolution urban DEMs. The experimental results demonstrate that MSM-CNN provides a powerful tool to reconstruct the topographic structures formed by road networks and building clusters from the low-resolution DEMs. Sufficient representation of these urban topographic structures are crucial for depicting urban hydrological processes, such as predicting surface runoffs and flooding with acceptable accuracy. Therefore, the MSM-CNN model and the reconstructed high-resolution DEM products can be used to support a range of applications in the water science and engineering fields, including urban flood risk management, and drainage systems planning and design. This is crucial for many cities in developing countries where high-resolution data are often scarce or even unavailable. It should be noted that the MSM-CNN model is not restricted by application to urban topography but is also applicable to the more natural topography in rural catchments to create high-resolution DEM data to support water resource management, natural hazard risk reduction, and many other forms of water engineering research and applications.

6. Conclusions

In this paper, we proposed an innovative deep machine learning approach to reconstruct high-resolution urban DEMs from low-resolution equivalents. In order to effectively account for the complexity of urban topography, a multi-scale CNN model was utilized to enhance the reconstruction quality. After the correlations between the low- and high-resolution urban DEMs are learned by the developed MSM-CNN model, an urban DEM at a specified high resolution can be accurately restored from a low-resolution dataset.

To evaluate the performance of MSM-CNN, a two-level accuracy assessment procedure involving both numerical accuracy and morphological accuracy was also designed and was used to compare the MSM-CNN with other DEM reconstruction methods including IDW, BI, CC, and KI. The results confirmed that MSM-CNN can effectively restore the high-resolution urban DEMs of 0.5 m from the low-resolution DEMs of 2, 4, and 8 m. The MSM-CNN products were also consistently better than those produced using alternative methods, in terms of visual assessment, and also numerical and morphological accuracy.

The promising results demonstrated that MSM-CNN provides a promising tool in generating high-resolution DEMs in cities from low-resolution DEMs, instead of surveying the whole region. In recent years, a number of global DEM products have been released to provide better resolution to represent urban topography, such as ALOS AW3D, NEXTMAP World 10, and WorldDEM. These open datasets can be explored and used to support the application of MSM-CNN to reconstruct high-resolution DEMs in cities across the world, which may potentially help address the challenging data scarcity issue and will have profound implications in many water-related applications, particularly in many of the developing countries.

Author Contributions: L.J. conceived and designed the methodology, wrote the original draft, and edited the manuscript; Y.H. wrote the CNN code, performed the CNN model training and testing, and contributed to the writing; X.X. contributed to the research ideas, supervised the experiments, and reviewed and edited the manuscript; Q.L. proposed the research ideas, discussed the structure of the paper, and reviewed and edited the manuscript; A.S. contributed to the research ideas, and reviewed and edited the manuscript; S.R.K. helped to revise the manuscript. All authors have read and approved the manuscript.

Funding: This work was supported by the U.K. Natural Environment Research Council (NERC) through the WeACT project (grant number NE/S005919/1), ValBGI project (grant number NE/S00288X/1) and Luanhe Living Lab project (grant number NE/S012427/1), the National Natural Science Foundation of China (grant numbers 41501445, 41701450, 41571398), State Major Project of Water Pollution Control and Management (grant number 2017ZX07603-001), China Postdoctoral Science Foundation (grant number 2018M642146), Jiangsu Planned Projects for Postdoctoral Research Funds (grant number 2018K144C), Anhui overseas visiting projects for outstanding young talents in Colleges and universities (grant number gxgwx2018078), and Key Project of Natural Science Research of Anhui Provincial Department of Education (grant number KJ2017A416).

Acknowledgments: The authors express their gratitude to the data support from Environment Agency (<https://environment.data.gov.uk/ds/survey/index.jsp#/survey>), and Digimap (<https://digimap.edina.ac.uk>). Many thanks are also given to reviewers and editors for providing constructive comments for this work.

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

References

1. Bishop, M.P.; James, L.A.; Shroder, J.F.; Walsh, S.J. Geospatial technologies and digital geomorphological mapping: Concepts, issues and research. *Geomorphology* **2012**, *137*, 5–26. [[CrossRef](#)]
2. Liu, X.; Tang, G.A.; Yang, J.; Shen, Z.; Pan, T. Simulating evolution of a loess gully head with cellular automata. *Chin. Geogr. Sci.* **2015**, *25*, 765–774. [[CrossRef](#)]
3. Mondal, A.; Khare, D.; Kundu, S.; Mukherjee, S.; Mukhopadhyay, A.; Mondal, S. Uncertainty of soil erosion modelling using open source high resolution and aggregated DEMs. *Geosci. Front.* **2017**, *8*, 425–436. [[CrossRef](#)]
4. Li, J.; Wong, D.W.S. Effects of DEM sources on hydrologic applications. *Comput. Environ. Urban* **2010**, *34*, 251–261. [[CrossRef](#)]

5. Moore, I.D.; Grayson, R.B.; Ladson, A.R. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrol. Process.* **1991**, *5*, 3–30. [[CrossRef](#)]
6. O’Loughlin, F.E.; Paiva, R.C.D.; Durand, M.; Alsdorf, D.E.; Bates, P.D. A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote Sens. Environ.* **2016**, *182*, 49–59. [[CrossRef](#)]
7. Ramirez, J.A.; Rajasekar, U.; Patel, D.P.; Coulthard, T.J.; Keiler, M. Flood modeling can make a difference: Disaster risk-reduction and resilience-building in urban areas. *Hydrol. Earth Syst. Sc. Discuss.* **2016**, 1–25. [[CrossRef](#)]
8. Leitão, J.P.; de Sousa, L.M. Towards the optimal fusion of high-resolution Digital Elevation Models for detailed urban flood assessment. *J. Hydrol.* **2018**, *561*, 651–661. [[CrossRef](#)]
9. Shan, J.; Aparajithan, S. Urban DEM generation from raw LiDAR data. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 217–226. [[CrossRef](#)]
10. Chen, Z.Y.; Xu, B.; Gao, B.B. An Image-Segmentation-Based Urban DTM Generation Method Using Airborne Lidar Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 496–506. [[CrossRef](#)]
11. Meesuk, V.; Vojinovic, Z.; Mynett, A.E.; Abdullah, A.F. Urban flood modelling combining top-view LiDAR data with ground-view SfM observations. *Adv. Water Resour.* **2015**, *75*, 105–117. [[CrossRef](#)]
12. Abdullah, A.F.; Vojinovic, Z.; Price, R.K.; Aziz, N.A.A. A methodology for processing raw LiDAR data to support urban flood modelling framework. *J. Hydroinform.* **2012**, *14*, 75–92. [[CrossRef](#)]
13. Tsubaki, R.; Fujita, I. Unstructured grid generation using LiDAR data for urban flood inundation modeling. *Hydrol. Process.* **2010**, *24*, 1404–1420. [[CrossRef](#)]
14. Abdullah, A.F.; Vojinovic, Z.; Price, R.K.; Rahman, A.A. Lidar Filtering Algorithms and DTM Generation for Urban Flood Modelling Applications: Review of Current Algorithms and Filters Test. In Proceedings of the 8th International Conference on Urban Drainage Modelling, Tokyo, Japan, 7–11 September 2009.
15. Wang, Y.; Chen, A.S.; Fu, G.; Djordjević, S.; Zhang, C.; Savić, D.A. An integrated framework for high-resolution urban flood modelling considering multiple information sources and urban features. *Environ. Modell. Softw.* **2018**, *107*, 85–95. [[CrossRef](#)]
16. Hawker, L.; Bates, P.; Neal, J.; Rougier, J. Perspectives on digital elevation model (DEM) simulation for flood modeling in the absence of a high-accuracy Open Access global DEM. *Front. Earth Sci.* **2018**, *6*, 223. [[CrossRef](#)]
17. Kulp, S.A.; Strauss, B.H. CoastalDEM: A global coastal digital elevation model improved from SRTM using a neural network. *Remote Sens. Environ.* **2018**, *206*, 231–239. [[CrossRef](#)]
18. Liu, Y.; Bates, P.D.; Neal, J.C.; Yamazaki, D. Bare-earth DEM Generation in Urban Areas Based on a Machine Learning Method. In Proceedings of the American Geophysical Union, Fall Meeting 2019, San Francisco, CA, USA, 9–13 December 2019.
19. Aguilar, F.J.; Agüera, F.; Aguilar, M.A.; Carvajal, F. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 805–816. [[CrossRef](#)]
20. Heritage, G.L.; Milan, D.J.; Large, A.R.; Fuller, I.C. Influence of survey strategy and interpolation model on DEM quality. *Geomorphology* **2009**, *112*, 334–344. [[CrossRef](#)]
21. Wise, S. Cross-validation as a means of investigating DEM interpolation error. *Comput. Geosci.* **2011**, *37*, 978–991. [[CrossRef](#)]
22. Arun, P.V. A comparative analysis of different DEM interpolation methods. *Egypt. J. Remote Sens. Space Sci.* **2013**, *16*, 133–139.
23. Tan, M.L.; Ramli, H.P.; Tam, T.H. Effect of DEM resolution, source, resampling technique and area threshold on SWAT outputs. *Water Resour. Manag.* **2018**, *32*, 4591–4606. [[CrossRef](#)]
24. Tran, T.A.; Raghavan, V.; Masumoto, S.; Vinayaraj, P.; Yonezawa, G. A geomorphology-based approach for digital elevation model fusion—case study in Danang city, Vietnam. *Earth Surf. Dynam.* **2014**, *2*, 403–417. [[CrossRef](#)]
25. Yue, L.; Shen, H.; Yuan, Q.; Zhang, L. Fusion of multi-scale DEMs using a regularized super-resolution method. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 2095–2120. [[CrossRef](#)]
26. Mason, D.C.; Trigg, M.; Garcia-Pintado, J.; Cloke, H.L.; Neal, J.C.; Bates, P.D. Improving the TanDEM-X Digital Elevation Model for flood modelling using flood extents from Synthetic Aperture Radar images. *Remote Sens. Environ.* **2016**, *173*, 15–28. [[CrossRef](#)]
27. Li, X.; Shen, H.; Feng, R.; Li, J.; Zhang, L. DEM generation from contours and a low-resolution DEM. *ISPRS J. Photogramm. Remote Sens.* **2017**, *134*, 135–147. [[CrossRef](#)]

28. Yue, L.; Shen, H.; Zhang, L.; Zheng, X.; Zhang, F.; Yuan, Q. High-quality seamless DEM generation blending SRTM-1, ASTER GDEM v2 and ICESat/GLAS observations. *ISPRS J. Photogramm. Remote Sens.* **2017**, *123*, 20–34. [[CrossRef](#)]
29. Xu, Z.; Wang, X.; Chen, Z.; Xiong, D.; Ding, M.; Hou, W. Nonlocal similarity based DEM super resolution. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 48–54. [[CrossRef](#)]
30. Chen, Z.; Wang, X.; Xu, Z. Convolutional neural network based DEM super resolution. In Proceedings of the XXIII ISPRS Congress, Prague, Czech Republic, 12–19 July 2016.
31. Moon, S.; Choi, H.L. Super Resolution Based on Deep Learning Technique for Constructing Digital Elevation Model. In Proceedings of the American Institute of Aeronautics and Astronautics SPACE Forum, Long Beach, CA, USA, 13–16 September 2016.
32. Liu, C.; Du, W.; Tian, X. Lunar DEM Super-Resolution Reconstruction via Sparse Representation. In Proceedings of the 2017 10th Image and Signal Processing, BioMedical Engineering and Informatics, Shanghai, China, 14–16 October 2017.
33. Xu, Z.; Chen, Z.; Yi, W.; Gui, Q.; Hou, W.; Ding, M. Deep gradient prior network for DEM super-resolution: Transfer learning from image to DEM. *ISPRS J. Photogramm. Remote Sens.* **2019**, *150*, 80–90. [[CrossRef](#)]
34. Mark, O.; Weesakul, S.; Apirumanekul, C.; Aroonnet, S.B.; Djordjević, S. Potential and limitations of 1D modelling of urban flooding. *J. Hydrol.* **2004**, *299*, 284–299. [[CrossRef](#)]
35. Ozdemir, H.; Sampson, C.C.; de Almeida, G.A.M.; Bates, P.D. Evaluating scale and roughness effects in urban flood modelling using terrestrial LIDAR data. *Hydrol. Earth Syst. Sc.* **2013**, *10*, 5903–5942. [[CrossRef](#)]
36. Leitão, J.P.; Moy de Vitry, M.; Scheidegger, A.; Rieckermann, J. Assessing the quality of digital elevation models obtained from mini unmanned aerial vehicles for overland flow modelling in urban areas. *Hydrol. Earth Syst. Sc.* **2016**, *20*, 1637–1653. [[CrossRef](#)]
37. Wang, C.; Yang, Q.; Jupp, D.L.B.; Pang, G. Modeling change of topographic spatial structures with DEM resolution using semi-variogram analysis and filter bank. *ISPRS Int. J. Geo Inf.* **2016**, *5*, 107. [[CrossRef](#)]
38. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
39. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)]
40. Abdel-Hamid, O.; Mohamed, A.R.; Jiang, H.; Deng, L.; Penn, G.; Yu, D. Convolutional neural networks for speech recognition. *IEEE Trans. Audio Speech Lang. Proc.* **2014**, *22*, 1533–1545. [[CrossRef](#)]
41. Chen, C.; Seff, A.; Kornhauser, A.; Xiao, J. Deepdriving: Learning Affordance for Direct Perception in Autonomous Driving. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.
42. Gu, J.; Wang, Z.H.; Kuen, J.; Ma, L.Y.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.X.; Wang, G.; Cai, J.F.; et al. Recent advances in convolutional neural networks. *Pattern Recogn* **2018**, *77*, 354–377. [[CrossRef](#)]
43. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Proceedings of the International Conference on Neural Information Processing Systems, Doha, Qatar, 12–15 November 2012.
44. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
45. 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, Las Vegas, NV, USA, 27–30 June 2016.
46. Hui, Z.; Wang, X.; Gao, X. Fast and accurate single image super-resolution via information distillation network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018.
47. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 8–10 June 2015.
48. Lee, C.Y.; Xie, S.; Gallagher, P.; Zhang, Z.; Tu, Z. Deeply-supervised nets. In Proceedings of the Artificial Intelligence and Statistics, San Diego, CA, USA, 9–12 May 2015.
49. Xie, S.; Tu, Z. Holistically-nested edge detection. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.

50. Lai, W.; Huang, J.; Ahuja, N.; Yang, M. Deep laplacian pyramid networks for fast and accurate super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 22–25 July 2017.
51. Kingma, D.P.; Adam, J.B. Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
52. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 13–16 December 2015.
53. Liu, K.; Song, C.; Ke, L.; Jiang, L.; Pan, Y.; Ma, R. Global open-access DEM performances in Earth's most rugged region High Mountain Asia: A multi-level assessment. *Geomorphology* **2019**, *338*, 16–26. [[CrossRef](#)]
54. Wilson, J.P. Digital terrain modeling. *Geomorphology* **2012**, *137*, 107–121. [[CrossRef](#)]
55. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [[CrossRef](#)]
56. James, M.R.; Robson, S. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surf. Proc. Land.* **2014**, *39*, 1413–1420. [[CrossRef](#)]
57. Gonçalves, J.A.; Henriques, R. UAV photogrammetry for topographic monitoring of coastal areas. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 101–111. [[CrossRef](#)]
58. Florinsky, I.V.; Kurkov, V.M.; Bliakharskii, D.P. Geomorphometry from unmanned aerial surveys. *T. GIS* **2018**, *22*, 58–81. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).