applied sciences

MDPI

Article

Deep Learning Based Urban Building Coverage Ratio Estimation Focusing on Rapid Urbanization Areas

Quang Hoai Le [1], Hyunkyu Shin [2,*], Nahyun Kwon [3,*], Jongnam Ho [1] and Yonghan Ahn [3]

[1] Department of Smart City Engineering, Hanyang University ERICA, Ansan 15588, Korea
[2] Center for AI Technology in Construction, Hanyang University ERICA, Ansan 15588, Korea
[3] Department of Architectural Engineering, Hanyang University ERICA, Ansan 15588, Korea
* Correspondence: hyunkew@hanyang.ac.kr (H.S.); nhkwon78@naver.com (N.K.)

Abstract: Urban parameters, such as building density and the building coverage ratio (BCR), play a crucial role in urban analysis and measurement. Although several approaches have been proposed for BCR estimations, a quick and effective tool is still required due to the limitations of statistical-based and manual mapping methods. Since a building footprint is crucial for the BCR calculation, we hypothesize that Deep Learning (DL) models can aid in the BCR computation, due to their proven automatic building footprint extraction capability. Thus, this study applies the DL framework in the ArcGIS software to the BCR calculation task and evaluates its efficiency for a new industrial district in South Korea. Although the accuracy achieved was limited due to poor-quality input data and issues with the training process, the result indicated that the DL-based approach is applicable for BCR measuring, which is a step toward suggesting an implication of this method. Overall, the potential utility of this proposed approach for the BCR measurement promises to be considerable.

Keywords: building coverage ratio; deep learning; urban management; urban density; mask R-CNN

Citation: Le, Q.H.; Shin, H.; Kwon, N.; Ho, J.; Ahn, Y. Deep Learning Based Urban Building Coverage Ratio Estimation Focusing on Rapid Urbanization Areas. *Appl. Sci.* **2022**, *12*, 11428. https://doi.org/10.3390/app122211428

Academic Editors: Wei-Ling Hsu, Yan-Chyuan Shiau and Hsin-Lung Liu

Received: 7 October 2022
Accepted: 7 November 2022
Published: 10 November 2022

1. Introduction

Urbanization is an inevitable part of development in countries all over the world. Expanding cities have led to a considerable increase in built-up areas [1], intensifying environmental degradation, global warming, and the urban heat island effect [2,3]. Buildings and urban structures also influence environmental conditions, a critical factor determining the quality of life in metropolitan areas [4]. High building density and a lack of control over built-up environment parameters, such as the building coverage ratio (BCR) and the floor area ratio (FAR), is exacerbating urbanization issues such as urban sprawl [5], the deprivation of the right to sunlight, and a lack of sufficient infrastructure provision [6]. Accurate insights into those parameters (e.g., BCR and FAR) and their relationship with the urban condition is crucial for urban analysis to mitigate the negative consequences of urbanization and, therefore, vital for urban status monitoring and analysis [7].

Understanding the BCR parameter contributes to better planning decisions as it plays a central role in a wide range of urban phenomena [8]. The BCR impacts ground surface reflectivity, airflow, and heat transmission in the atmosphere over metropolitan areas, causing Urban Heat Islands (UHI) [9]. Scholars indicate that controlling the BCR may help reduce this phenomenon [8,10]. Moreover, choosing an appropriate BCR for urban planning and building design may be an effective strategy to address difficulties created by urbanization, and to enhance both environmental conditions and citizens' quality of life. For instance, in the context of urban management, the maximum (or "legal") BCR and FAR limits are fundamental tools for district-scale zoning regulation that legally control the shape and volume of buildings, supporting more effective urban planning as well as improving city aesthetics [11]. However, BCR violations, such as encroachment construction, have become a major problem for urban management in several areas [12].

Although an effective BCR monitoring tool will support the efforts of city planners to enforce existing urban management and planning regulations [13], it has yet to be achieved. Due to the tendency toward urban sprawl, cities are dynamic, continually expanding, and changing dramatically, which is especially the case in developing countries [14,15]. This makes extracting urban features such as the BCR even more challenging, and many urban communities lack access to a comprehensive BCR dataset for their area [16].

Recent advances in Deep Learning (DL) have had a significant impact on fields such as spatial analysis and remote sensing. DL is a type of Machine Learning (ML) that uses multiple layers of neurons with complicated structures, or applies non-linear changes to high-level model abstractions of the original input [17]. Thanks to DL, extracting urban features from high-resolution spatial images is now becoming more accessible and faster [18–20]. Several state-of-the-art detection models are now available that deliver superior performance for building footprint extraction [21–23]. The building footprint can be used to describe the share of the building area in the total block area, which is crucial in the BCR measuring [24]. Therefore, automated building footprint extraction based on DL has the potential to speed up the BCR calculation process considerably, compared to manual mapping methods. In the literature review, we summarize the advantages and shortfalls of proposed BCR calculation methods. Due to the limitations of conventional and statistical-based approaches, the demand for a quick and robust BCR measuring approach was proven. Most previous BCR calculation methods have required a pre-processing phase to create the desired input, after which the calculation process can proceed. This makes the process both protracted and complex. The use of DL should make this pre-processing step simpler, or even unnecessary, due to the automatic detecting and generating of the building footprint outlines. In this way, users may concentrate on evaluating and comprehending the pattern rather than pre-processing data. We hypothesize that Deep Learning (DL) models can help with the BCR computing due to their proven automatic building footprint extraction capability. Thus, this study applies DL to a BCR calculating task to test the performance of this approach and explore the correctness achieved and the factors that affect the accuracy.

In this study, an alternative BCR calculation approach based on the DL workflow in the ArcGIS software is presented and tested by utilizing a DL-based method to measure and visualize the BCR distribution in sections of the Sihwa industrial district in Ansan City, South Korea. The main goal is to identify the feasibility of DL in measuring the BCR and propose a framework to apply this approach in urban management. The findings of this study could accelerate the use of DL in measuring the BCR to support city managers and planners striving to cope with rapid urbanization.

2. Literature Review

As the review demonstrates in Table 1, a great deal of effort has been devoted to finding a good way to measure urban density and building coverage in recent decades. In order to shed light on the relationship between the BCR and land-use change, Magayanti et al. [12] carried out a field study to collect evidence. However, site surveys are labor-intensive, time-consuming, and have limited utility. As the drawbacks of census and site survey methods are significant, BCR datasets generated by sources such as satellite and airborne imagery could be attractive alternatives. In a study that measured variations in building density by using the BCR and the FAR indices, Pan et al. [25] proposed a statistical method to measure the BCR from satellite imagery using ArcMap and ENVI to obtain building heights and coverage areas. For this statistical analysis, building contour retracing was manually repeated to compute areas. Although this building information extraction technique was challenging or incorrect when dealing with complex building geometries, this method gives a quantitative strategy to monitoring building density distribution.

In response to manual mapping challenges, researchers have developed an automatic approach to extract urban density measurements from polarimetric synthetic aperture radar (SAR) imagery. Kajimoto and Susaki [26] proposed novel urban compactness estimating

method based on a polarization orientation angle (POA) algorithm. However, a subsequent study by Susaki et al. [27] pointed out a number of issues with this method that limited the utility of the urban density maps generated that they addressed by extending the POA algorithm to develop an alternative method for estimating urban density from SAR images. Other research has focused on using digital surface model (DSM) and LiDAR data. For example, Gonzalez-Aguilera et al. [28] proposed an automated approach to assess building geometry information. They used a progressive triangular irregular network (TIN) to detect buildings and generate building outlines to extract parameters, such as the BCR and the FAR. Although LiDAR data, satellite, and SAR imagery are all good sources for extracting urban features, they are not always available or affordable. Low-cost alternatives, such as Open Street Map (OSM), have also been pressed into service to analyze building compactness [29]. However, open data sources and OSM data vary from region to region, especially in developing countries, making the widespread use of OSM data for urban analysis somewhat problematic. Unlike satellite imagery, airborne images are much more accessible. For example, Emmanuel et al. [30] proposed a method for computing the BCR and the FAR from airborne images and DSM based on the mean-shift algorithm. The effectiveness of this approach is significant due to its utility in cases where only limited information or low spatial resolution data are available. However, the mean-shift-based method is less suitable for congested areas since buildings are detected as homogeneous objects, reducing model accuracy. By considering the importance of the geographic distribution of building coverage for sustainable development, Soliman et al. [16] developed a model for estimating building coverage and predicted the BCR distribution across the contiguous United States. This study provided quantitative information on the geographical distribution of building coverage at the national scale and may thus only be appropriate for large area urban analyses.

Table 1. Summary of Previous Studies.

Study	Objectives	Method	Remarks	Data	Scale
Pan et al., 2008 [25]	Measure the variation of building density by using BCR and FAR indices	Statistics, Quantitative	Inaccurate for complex buildings	SI	DS, BS
Kajimoto & Susaki, 2013 [26]	Build an algorithm to estimate urban density from Polar images	POA Algorithm	The urban density of multi-images generated is not comparable	SAR	RS
Gonzalez-Aguilera et al., 2013 [28]	Develop an automated approach to assess building geometric information (BCR, FAR)	LiDAR-derived building models	High cost associated with acquiring photos	DSM	BS, PS
Hecht et al., 2013 [29]	Describe methods to analyze building compactness	Quantitative	The completeness of OSM data varies from regions	OSM	BS, PS
Susaki et al., 2014 [27]	Propose an algorithm for estimating urban density from SAR image	Extended POA Algorithm	Suitable for large area	SAR	RS

Table 1. *Cont.*

Study	Objectives	Method	Remarks	Data	Scale
Emmanuel et al., 2016 [30]	Proposed an automated method for computing BCR and FAR	Modified Mean Shift Algorithm	Extracts buildings as homogeneous objects rather than distinguishing them	UAV, DSM	BS
Megayanti et al., 2018 [12]	Examine the extent to which implementation of BCR influences land-use changes	Field Survey, Quantitative	Time-consuming, intensive	Observation Data	BS, PS
Soliman et al., 2018 [16]	Develop a model for estimating building coverage	Statistics	Required various types of input data and complexity	Census data	RS

Note: RC: Regional-scale; DS: District-scale; BS: Block-scale; PS: Parcel-scale. SI: Satellite imagery, SAR: Synthetic Aperture Radar image, DSM: Digital model surface, OSM: Open Street map, UAV: Unmanned aerial vehicle (Drone).

There has recently been a shift of the BCR extracting method from low accessibility data sources to more accessible ones, partly due to easy access to new technologies, such as drones. Drone imagery has proved its quality and accessibility for remote sensing applications compared to satellite imagery [31,32]. The BCR extraction methods based on drone imagery can overcome the problems that plague satellite data sources, namely acquisition cost and accessibility. Besides, simplifying the pre-processing and calculation process would encourage the wider application of BCR calculating methods for real-world urban density analyses. Since previous studies required considerable human effort to be devoted to digitalizing mapping or pre-processing data, this was likely to result in lengthy analyses with high error rates when assessing large and complex data sets. Most previous statistical models are not well adapted to deal with the rapid growth of today's cities, failing to take advantage of the advent of urban Big Data [33]. The need for a robust and reliable BCR measuring tool to overcome the challenge of rapid urbanization is clear.

Thanks to the advancements in ML and DL algorithms, robust processing tools that can reliably and effectively extract the critical information contained in remote sensing data have been proposed. Leveraging the advantages of these new tools, several studies on urban analysis based on ML and DL achieved considerable results [22,34,35]. However, the research on utilizing DL in the BCR measuring is limited. The use of DL in BCR calculating has yet to be nurtured in previous studies. Therefore, this study, which presents and evaluates the use of DL for the BCR extraction from UAV imagery, is designed to take advantage of the new tools and resources that are becoming available to researchers.

3. Materials and Methods

The analysis was conducted in the 1st phase development of the Sihwa Industrial District, a National Industrial Complex located in Ansan-si, Gyeonggi-do, South Korea. The area is covered with multi types of building such as industrial buildings, detached and semi-detached buildings. The Korean National Geographic Information Service supplied the researchers with a collection of orthophotos taken by drones in 2016. These photos were 0.25 m resolution and utilized a coordinate system based on the GSR80 geodetic reference system. Orthophotos were used for this study because these images suitable for area computing. In contrast to an uncorrected aerial image, an orthophoto may be used to accurately measure distances since it is an exact portrayal of the earth's surface that has been corrected for topographic relief, lens distortion, and camera tilt [36]. In other words,

orthophotos have been geometrically rectified to have a consistent scale, which is crucial for achieving an accurate building area calculation.

The dataset includes images taken of two Zones, A and B (Figure 1), which were used to export image chips to form the training dataset. Due to a lack of available data for training deep learning models, images taken in Zone B were divided into two parts; one part was used to combine with that of Zone A to form the training dataset. This makes sure that the training data set is large enough and includes various building types for effective model training. The remaining part of Zone B was used to test model performance, extract the building footprints, and perform the BCR calculations.

Figure 1. Zone A and Zone B (within the red boundaries) of the Sihwa industrial complex.

3.1. Experimental Procedure

To test the efficacy of using the DL model to facilitate the BCR calculation process, the following three steps were used (Figure 2):

1. Establish the DL model. A backbone and input picture size that best suits the building detection task is specified to fine-tune the DL model to be used to extract building outlines from drone images.
2. Extract the building footprints. This step is to extract building footprints from the images captured in Zone B to produce the building footprint layer, the input for the next step.
3. BCR calculation and validation. The BCR can now be calculated at two levels, at the parcel scale (where BCR was measured based on buildings' area within a single land boundary) and block scale (where the boundary was determined by surrounding roads). Finally, the accuracy of the DL-based methods is evaluated.

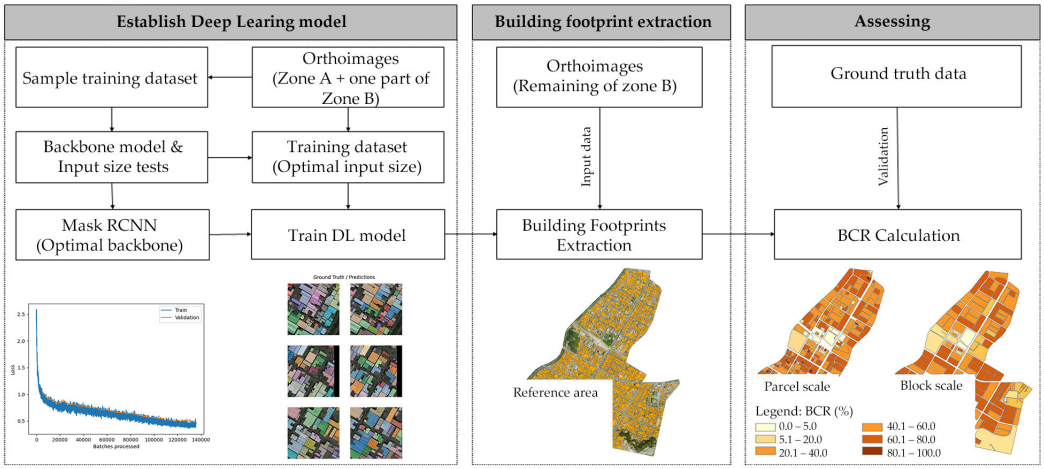


Figure 2. The experimental procedure.

3.2. Establish the DL Model

Mask R-CNN, introduced in 2017, is currently the most potent instance segmentation technique known [37]. It is widely used for instance segmentation, which is a precise delineation of objects in an image. Previous DL models, such as Single Shot Detector, RetinaNet, and YOLO, can recognize buildings [38,39]; they all produce a bounding box to show the location of a building on an image, which is insufficient for a BCR calculation. For the BCR calculations, a pixel-based output mask is necessary to identify building boundary so that the building area can be computed, which is fit with the ability of the Mask RCNN model. Therefore, it was used to extract the building footprint in this study. Mask RCNN is an extension of Faster R-CNN that predicts segmentation masks for each Region of Interest (RoI). The architecture of Faster R-CNN consists of two networks, backbone (ResNet, Inception, etc.) and the region proposal network. Region proposals are regions that contain the object on the feature map. Using either the RoI pool or the RoIAlign method, the proposed region sizes are fixed. The output of the RoIAlign layer is then fed to the mask head, which is comprised of two convolution layers. It generates a mask for each RoI, thereby pixel-by-pixel segmenting an image (Figure 3).

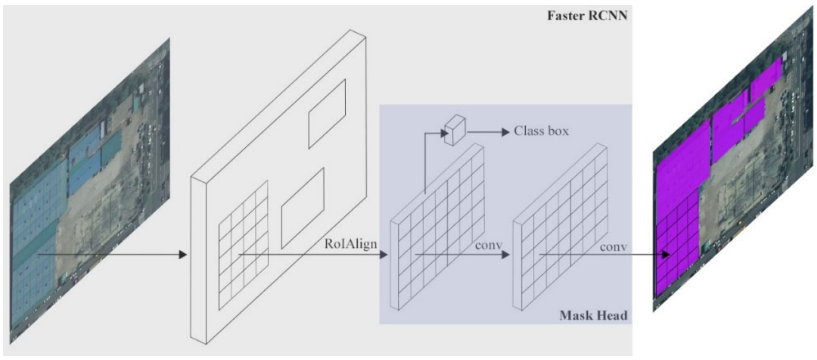
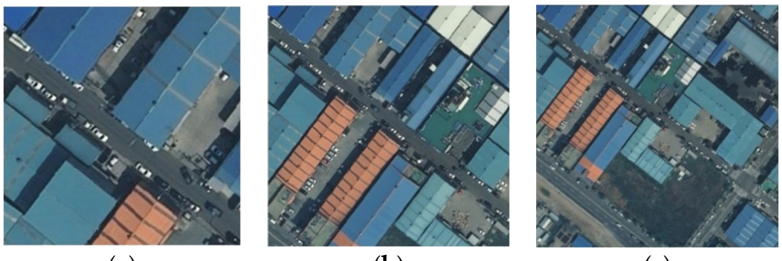


Figure 3. The mask RCNN framework for instance segmentation.

Using a ResNet backbone for feature extraction improves accuracy and speed when used with Mask R-CNN [40]. There are several versions of ResNet, depending on the number of layers contained (either 18, 34, 50, 101, or 152); the ResNet versions containing 18 and 34 layers have lower accuracy than those with more layers [41]. Therefore, three types of residual networks (those containing 50, 101, and 152 layers) are tested in this study. Besides, the size of the training picture affects the accuracy of localization and the contextual information it can receive [42]. Choosing the right size for the training photos is critical in the building detection process. The best picture size was chosen by comparing model performance with different training image sizes, as shown in Figure 4.



(**a**) (**b**) (**c**)

Figure 4. The receptive field of input image sizes: (**a**) 256 × 256, (**b**) 512 × 512, and (**c**) 768 × 768 pixels.

In order to choose an optimal backbone model and a suitable size of training data for the DL model in the building footprint extraction task, the performance of Mask R-CNN with various backbones and training picture sizes was evaluated. Three residual networks were trained on three data sets to find the optimal combination. Each training dataset has

4402 picture chips of 256, 512, and 768 pixels each. Popular accuracy measurement methods for object detection and semantic segmentation, namely *Precision*, *Recall*, *F1-score*, and *Average Precision* (AP), were used to measure the model performance [20]. Mask R-CNN provides both a bounding box and mask. Precision recall values were calculated from the bounding box output. Generally, if the Intersection over Union (IoU) value is greater than 0.5, the predicted bound box is True; otherwise, the output is False. IoU metric in object detection evaluates the degree of ground truth and prediction overlap (Equation (1)). The results extracted from the bounding box are calculated as the precision-recall value, where *Precision* is the ratio of the number of true positives to the total number of predictions (Equation (2)) and *Recall* is defined as the ratio of the number of true positives to the total number of positive predictions (Equation (3)). The *F1 Score* is the weighted average of the precision and the recall, and the values range from 0 to 1, with 1 denoting the highest level of precision (Equation (4)). A precision-recall curve is summarized by *Average Precision* as the weighted mean of precisions at each threshold, with the increase in recall from the previous threshold serving as the weight (Equation (5)).

$$IoU = \frac{\text{Area}(\text{ground} \cap \text{predict})}{\text{Area}\ (\text{ground} \cup \text{predict})} \tag{1}$$

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$F1\ score = \frac{\text{P} \times \text{R}}{(\text{P} + \text{R})/2} \tag{4}$$

$$AP = \sum_{n} (R_n - R_{n-1}) P_n \tag{5}$$

where TP is the number of True positives; FP is the number of False positives; and FN is the number of False negatives; P is *Precision*; R is *Recall*, and AP is *Average Precision*.

Table 2 shows the correlation between the number of convolutional layers, training image size, and training time. The performance of the Mask R-CNN model varied for different training image sizes and backbone models. Doubling the image size from 256 × 256 to 512 × 512 pixels increased the model performance in three ResNet models, but tripling the size to 768 × 768 pixels degraded the performance in both ResNet101 and ResNet152. There was an increase in performance when the network was deeper, increasing from 50 to 101 layers deep. However, when the depth increased to 152 layers, ResNet152 delivered a lower accuracy than ReNet101. Although the number of stacked layers might enhance the model's features, a deeper network can reveal any degradation issues. This is because as the number of layers in a neural network grows, the accuracy levels may become saturated and thus gradually decline, thus degrading the model's performance considerably. With the same backbone model (ResNet101), the training time with input image sizes, at 768 and 512 pixels, was 19 h 35 min, and 9 h 56 min, respectively; the training image size clearly affects the total training time considerably. Although the latter has slightly lower accuracy than the former, it performed better in terms of AP, F1, and Recall, at 0.3976, 0.5980, and 0.4668, respectively. The pattern of training times was similar when increasing layers in the ResNet model, with deeper layers leading to longer learning processes. Although ResNet50 needs less training time than others, the performance is lower for all four accuracy metrics. Based on these test results, ResNet101 with a training image size of 512 pixels, seems a suitable combination as it delivers the optimum results, therefore, it was used for model fine-tuning in the next step.

Table 2. Comparing the performance of Mask RCNN with ResNet models and input sizes. Here, epochs = 30, batch size = 2, and 10% of the 4402 training samples were used for validation.

Backbone	Input Size	Precision	Recall	F1_Score	AP	Training Time
ResNet 50	256	0.5353	0.3163	0.3977	0.1693	3 h 51 m
	512	0.7192	0.3571	0.4773	0.2641	8 h 33 m
	768	**0.8360**	**0.3597**	**0.5030**	**0.3008**	16 h 53 m
ResNet 101	256	0.6137	0.3648	0.4576	0.2239	4 h 39 m
	512	0.8318	**0.4668**	**0.5980**	**0.3976**	9 h 56 m
	768	**0.8534**	0.4405	0.5810	0.3808	19 h 35 m
ResNet 152	256	0.6573	0.4175	0.5107	0.2748	5 h 21 m
	512	0.7980	**0.4065**	**0.5386**	**0.3244**	11 h 33 m
	768	**0.8216**	0.3759	0.5156	0.3128	23 h 50 m

3.3. Training Model and Building Footprint Extraction

Mask R-CNN, with ResNet101 as a backbone, was trained with a bigger dataset in this step to increase the model accuracy. Due to the lack of a public training dataset for the study area, 8201 building footprints were created manually and labeled as building in the ArcGIS Pro 2.8 to prepare data for the training process. The orthophotos were used to export image chips using the geoprocessing tool Export Training Data for Deep Learning in ArcGIS. The rotation angle was set at 90 degrees for data augmentation. The training dataset consisted of 23,217 image chips, each with a size of 512 × 512 pixels, processed on a Windows 10, Deep Learning workflow in ArcGIS with the following computer specification: GPU NVIDIA GTX 2080Ti, 32GB RAM. As noted above, the experimental model selected was Mask R-CNN with ResNet101 as the backbone, epochs = 20, batch size = 2, and the framework was Pytorch.

The performance of the Mask R-CNN model was measured using three of the accuracy metrics used in the previous step, namely Recall, F1-Score, Average Precision. However, there was no major increase in the trained model's performance, largely due to the variation in buildings' characteristics in different areas, as missing building features leads to low model performance. Thus, in the next step, we used this trained model as a pre-trained model to fine-tune a new model with a dataset consisting of part of Zone B, including 3135 image chips. This step was designed to update the model by enabling it to learn building features that were not present in the Zone A dataset and thus improve the model accuracy. The results indicate that the model performance significantly improved after training with the previously established pre-trained model. Table 3 compares the performance achieved by the trained models after the first and second training step. Although the precision of the model after the second training step was lower than that after the initial training, the scores for Recall, F1 score, and average accuracy are all considerably higher for the model that had undergone the second training step.

Table 3. Evaluating the accuracy of the DL model.

	Precision	Recall	F1 Score	AP
1st Training	**0.8080**	0.5718	0.67697	0.4664
2nd Training	0.7646	**0.7493**	**0.7569**	**0.5829**

The fine-tuned DL model was then used to detect building footprints. The geoprocessing tool in ArcGIS, Detect Object Using Deep Learning, generated the polygon feature class showing the building footprint for the BCR calculations. All the buildings in the test area of over 10 square kilometers were detected in less than 5 min, indicating the speed

of DL for building footprint extraction. The output was a layer of building footprints identified by the building footprint detection process. However, as Mask R-CNN generates its results by segmenting an image in a pixel-to-pixel manner, the edges of the detected building shape are not straight. This was addressed by utilizing the Regularize Building Footprints tool to post-process the detected building footprints to normalize the footprint of the constructing polygons.

3.4. BCR Calculation

In this step, the extracted building footprints from the previous step were used to compute coverage ratios. Tabular data from the extracted building footprint layers were used for calculations. The BCR is defined as the percentage of the lot area covered by buildings (Equation (6)). The BCR was calculated at both the block scale and the parcel scale, with the site coverage including all the building footprints inside a block or a parcel. An assessment of 116 blocks and 267 parcels in Zone B was conducted to evaluate the accuracy of the BCR calculation based on the DL model.

$$BCR(\%) = \frac{\mathrm{Sb}}{\mathrm{Sl}} \tag{6}$$

$$Accuracy(\%) = \frac{|eBCR - aBCR|}{aBCR} \tag{7}$$

where Sb is the area covered by a building, defined by the building footprint, and Sl is the area of the site where the building is located.

In order to calculate the BCR, the building coverage area and block areas are required. Thanks to DL's building footprint extraction ability, it can facilitate the BCR estimation. The DL is appropriate for extracting building areas; however, it may not be appropriate for extracting parcel areas. The building's area is dependent on construction, and it can change time by time due to the expansion demand, while the parcels' area is fixed. Therefore, extracting building areas based on DL models is more reasonable than extracting parcel areas. Developing a DL model takes time and effort; thus, using the DL model to extract a constant variable like parcel area is not appropriate for DL nature. Besides, there are several sources of data that are used to acquire parcel areas more attainable than using DL, such as cadastral data, land development plans, etc. Therefore, in this study, we used cadastral data to extract parcel areas. The resulting BCR values for the blocks and parcels were then compared with the ground truth data to evaluate the model accuracy. Accuracy is the fit between extracted BCR (*eBCR*) and the actual BCR (*aBCR*) as shown in Equation (7). The low difference in the ratio between extracted BCR and the actual BCR indicates the high accuracy of the DL-based approach. This is the key point that we used to evaluate the effectiveness of the DL-based method in computing the BCR.

4. Results and Discussion

4.1. Results

The fine-tuned model can precisely generate the building footprints for buildings in the reference area. The comparison between segmentation results and the ground truth label indicated the model can also distinguish buildings that are close to each other by generating masks in different colors (Figure 5a). Since each building is identified as a distinct entity, the BCR calculation in dense areas is more accurate. In Figure 5b, results after the application of the fine-tuned model to generate building footprints from drone images taken in Zone B, used for the BCR calculation, are shown.

Figure 5. (**a**) The actual segmentation results plotted against the ground truth label, and (**b**) regularized building footprints overlaid on top of a sample area from validation dataset.

The BCR was calculated for each parcel and block in the referenced area based on the extracted building footprints. Then, the result was compared with the actual BCRs from the ground truth dataset to validate the accuracy of DL-based method. In order to identify the correlation between the two variables, the BCRs at both the block scale and parcel scale were plotted on scatter charts to highlight the relationship (Figure 6a,b). The R^2 values off 0.5899 and 0.7284, respectively, for the block scale and the parcel scale, although there are some outlier points where a calculated BCR value is either much larger or much smaller than that for the actual data. The outliers in the BCRs at the block scale were significantly greater than those at the parcel scale, leading to a considerably lower R^2 value. The outliers were mainly located in the low-density blocks with building densities below 20%; the fitting for the dense blocks was generally much better. Unlike the block scale, the pattern in the parcel scale data was less clear, with outliers appearing at many levels of the BCR, including in the range from 20% to 60%. Displaying the results as histogram charts (Figure 6c,d) revealed that a large number of samples had accuracies of over 85%. At both scales, the accuracy from 90% to 100% accounted for the most significant proportion of the results. The number of samples with an accuracy lower than 50% was considerable at the parcel scale, where 22 samples accounted for 8.24% of the total, while at the block scale just seven samples were below 50%, accounting for 6.04% of the total.

The accuracy of DL-based methods is shown in Table 4. At the block scale, the presence of even a few outliers reduced the average accuracy to 86.55%, even though the median was 94.07%. The median, which is the middle value in a BCR accuracy list ordered from smallest to largest, indicates that half of the samples were over 94.07%. The gap between average accuracy and median accuracy at the block scale was narrower than at the parcel scale, which suggests that the performance of our proposed DL-based method was more stable at a larger scale, where a lower level of detail is required.

For the purpose of gaining insight into the status of the building coverage ratio in the referenced area, the extracted BCR based on DL for the block and parcel scales is shown in scatter plots (Figure 7). Here, the coverage ratios between 60% and 70% accounted for a large proportion at the block scale, indicating that the study area is quite congested. The proportion of blocks with a coverage ratio lower than 40% was relatively small, and just a few blocks had BCR values higher than 75%. The coverage ratio distribution at the parcel scale was wider, ranging from 50% to 80%. Besides, a quantitative and visual view of the BCR distribution is achieved by using a thematic map, as shown in Figure 7c,d. The large blocks, which are usually parks, have lower levels of BCR than small blocks. Due to the construction of buildings, the coverage ratio is usually high in parcels that have a small area. Some of them have a coverage ratio of up to 80%, which is reaching the maximum legal BCR. This reflects the status of construction in a crowded industrial zone like Sihwa.

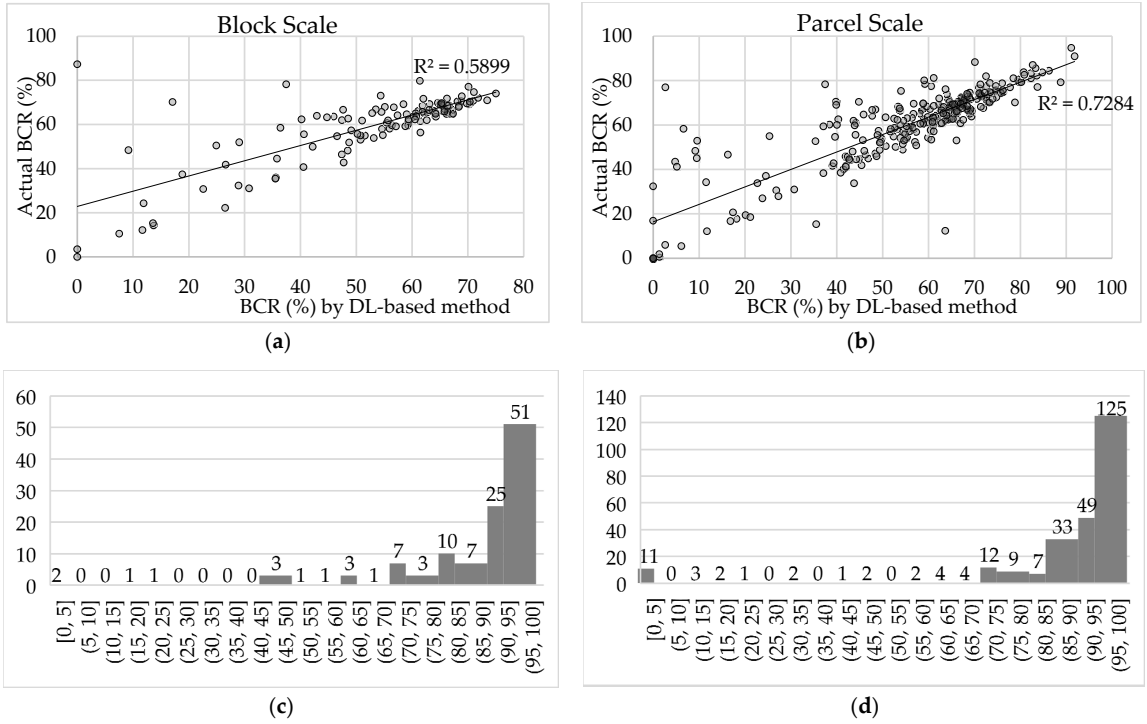


Figure 6. The fitting effect between BCR calculated using the DL-based method and the ground truth data (**a**) at the block scale, and (**b**) at the parcel scale. Accuracy distribution at (**c**) the block scale, and (**d**) the parcel scale.

Table 4. The accuracy of BCR by DL.

Scale	Block Scale		Parcel Scale	
Accuracy	Average	Median	Average	Median
Percentage (%)	86.55	94.07	85.21	94.29

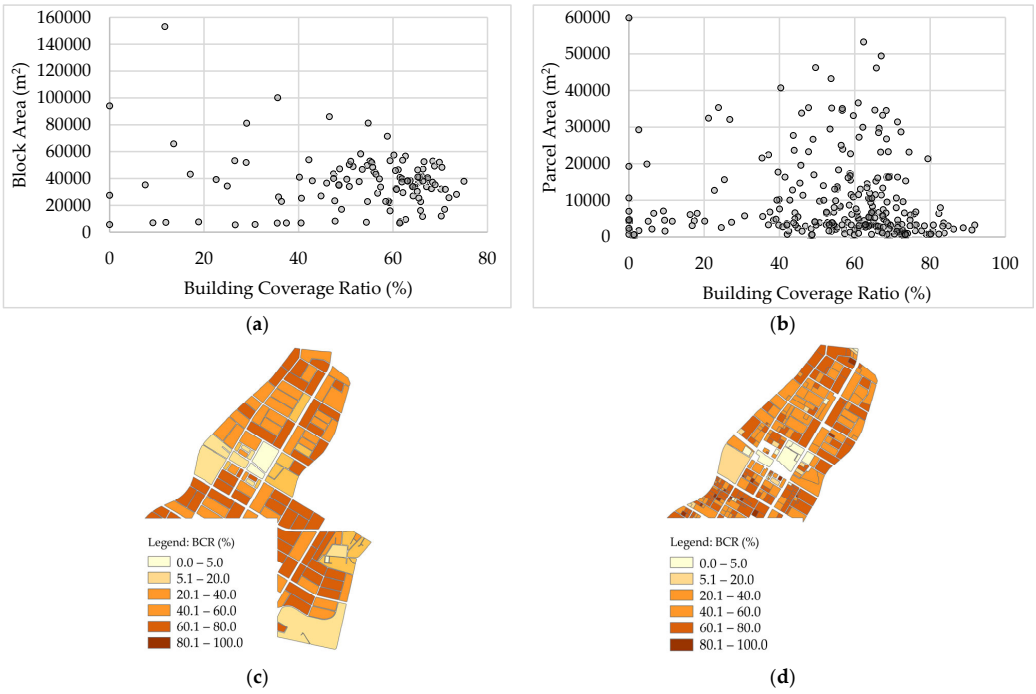


Figure 7. BCR distribution (**a**) at the block scale, and (**b**) at the parcel scale. Thematic map visualizing the BCR (**c**) at the block scale, and (**d**) at the parcel scale.

4.2. Discussion

In the rapid urbanization context, cities are continuously expanding, leading to difficulty for urban planners to evaluate the development of built-up areas and their parameters, such as the BCR. The proposed DL-based approach to measuring the BCR has been tested and evaluated through an assessment to monitoring the coverage status of urban areas at different scales. Through the accuracy assessment of the DL-based method in the BCR calculation, the potential of this approach has been revealed. Utilizing DL for the building footprint generation speeds up the BCR measurement and prevents errors and effort for manual mapping. Moreover, by displaying the coverage ratio of each block, a complete and quantitative image of the shape and spatial organization of the research area's urban environment was generated. Consequently, congested areas could be visualized easily, allowing decision-makers to identify over-construction blocks. This means that by comparing the extracted BCR with the legal BCR, tracking legal BCR compliance is possible. By indicating potential over-construction regions, planners can consider suitable redevelopment plans to control the BCR in regions such as industrial zones, business parks, and residential areas. Urban density management, therefore, is more effective and proactive.

Since drones can acquire aerial images quickly and at a lower cost than satellite imagery, the BCR extraction from drone images based on DL could be more accessible for decision-makers charged with carrying out urban density assessments. Suppose, for example, that the accuracy of the DL-based method can be improved. In that case, aerial inspection using a DL-based method could potentially replace fieldwork for measuring the coverage ratio of built-up areas in today's rapid urbanization context, which requires frequent updates to track our ever-changing urban environments. The BCR calculation based on DL has the potential to become a vital tool for urban management to support urban planning decisions. This tool will allow policymakers to perform city-wide building coverage investigations to identify over-developed regions and select low-density areas for upgrading. The proposed application of the DL-based method in urban building coverage survey is shown in Figure 8. The BCR status of certain areas can be measured by utilizing the proposed DL-base method. Then, the extracted BCR status can be visualized by using data visualization tools to monitor the BCR. Consequently, decision-makers can check whether the referenced areas exceed the legal BCR so that further actions can be considered. Besides, the BCR and the FAR can be used for urban density analysis. This provides a foundation for population estimation and infrastructure development plans.

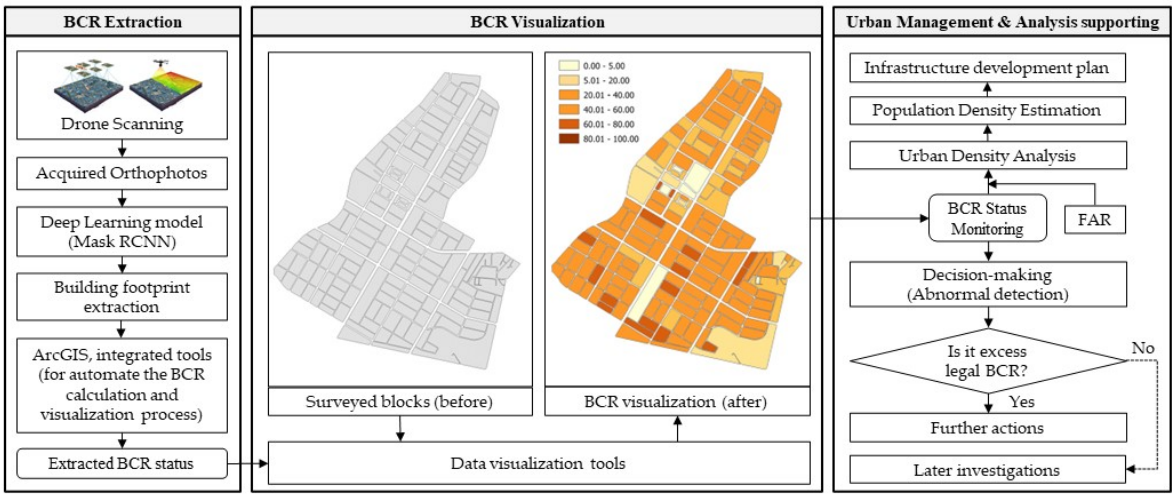


Figure 8. Proposed applications of the new BCR calculation approach based on DL for urban management and urban density analyses.

In countries that enforce urban regulations to control legal BCRs (e.g., Japan, Korea, and Vietnam), compliance tracking is crucial for urban management to prevent over-

construction. However, it is mostly carried out as part of the application process for building permits and completion. This means that encroachment construction that takes place after a building enters operation might be missed, leading to a BCR violation [13]. The BCR tracking after construction through site surveys is often difficult due to land accessibility and being time-consuming and costly. In light of our proposed DL-based method and drone mobility, acquiring blocks and parcels' BCR seems easier than conventional methods in terms of time, labor, and cost. Thus, monitoring the BCR at a city scale is possible to track legal BCR compliance, contributing to improving urban management. The use of aerial inspections to extract the BCR in urban areas might fast become a viable method for urban density analysis and management.

However, there are some obstacles to the proposed DL-based approach that need to be considered. Even though the DL-based approach could be useful for the BCR measuring, preparing a suitable training dataset is not straightforward [43]. Creating a DL model from scratch is time-consuming and requires specialized skills. This method might be suitable for those who already have a well-trained DL model, or they may choose to utilize a technique that produces training samples automatically [44]. Transfer learning could be another way to avoid the need to develop site-specific training data; this approach would allow users to save resources and improve efficiency when training a new model.

Besides, the performance of the DL model is also crucial in the BCR extraction. Basically, the degradation in model prediction leading to a decrease in the BCR measuring accuracy. In specific BCR blocks, or lots where the building is significantly modest and blends with the surroundings, the DL model is unable to generate the building footprint, leading to a significant issue in the BCR measuring. Additionally, a few building typologies might not be included in the training dataset, which can again lead to missing detection. Each type of building has a different profile; for instance, factories look distinctly different from other building typologies [45]. This means that a well-trained model that detects residential buildings with high accuracy may not turn in a good performance for other building types. Obviously, the training model plays a crucial role in building extraction; therefore, the training dataset should include as many different building typologies as possible to achieve a good result and in order to optimize the building detection process.

In addition, the building height results in the distortion effect, which also affects the accuracy of the BCR measurements. While a conventional building detection task simply requires the segmentation of buildings located on a spatial image, the information needed for the BCR calculations is more challenging. Although orthodox imagery was used for the current study to avoid distortion effect, there remains a problem. Namely, in some cases when generating the footprints for high-rise buildings, the DL model generated the building footprints that were bigger than the ground truth shapes, due to vertical structures being included. Consequently, the BCR values extracted from these buildings tend to be higher than is actually the case. In other words, to compute the correct BCR, a building's segmentation must be similar to its footprint. However, the distortion effect is increased in buildings that have high heights, leading to the visibility of buildings that include vertical structures rather than just the roofs. This might be caused by the limitations of precautions during orthoimage generation [36] causing issues in the building footprint generation for the BCR calculation. Further research should consider this issue to find a way to deliver more reliable results.

5. Conclusions

The advancements of ML and DL accelerate the proposal of robust urban assessment tools. However, utilizing DL algorithms to optimize the BCR calculation has yet to be focused upon. Therefore, this paper suggests a BCR calculation method based on the deep learning (DL) workflow in ArcGIS. The accuracy achieved by extracting the BCR from drone images based on DL was evaluated by analyzing the BCR values of 116 blocks and 267 parcels. This method is quicker and more effective than conventional methods, and it also overcomes the limitations of the previously proposed methods in terms of

distinguishing buildings and preprocessing data. The building footprints were extracted automatically, resulting in a considerably reduced process duration. Furthermore, the instance segmentation algorithm identified each building as a distinct entity, so that the BCR measured was more accurate, especially in dense areas. The results indicate that the accuracy of the DL-based method was considerable, with average and median accuracies at the block scale of 86.55% and 94.07%, and at the parcel scale of 85.21% and 94.29%, respectively. However, the accuracy of this approach relies heavily on the DL performance, which is affected by both the quality of the input data and the training process.

Based on the accuracy evaluation, we conclude that the DL-based approach is suitable for BCR measuring, which is a stepping stone to proposing an implication of this method. Techniques based on deep learning may aid in the monitoring of the BCR in urban management due to the robustness of DL and drone mobility. This approach has the potential to assist policymakers to analyze a city's building coverage and identify legal BCR compliance, thereby contributing to the prevention of encroachment construction and the control of urban density distribution. In equatorial or warm climatic regions where the urban heat island effect is more severe, urban density control is a workable strategy to mitigate negative effect this urbanization phenomenon.

The main limitations of this study are as follows. First, due to the lack of public datasets, the model was trained with a mixed-training dataset, which is not an ideal condition. The training dataset establishment needs to be improved so that better proven the generalization abilities of the model. Second, our proposed DL-based method was assessed for industrial zones, where the building type is uniform and lacks building-height diversity. Further research should consider the issue when assessing high buildings to find a way to deliver more reliable results. Besides, the cost-benefit of this method also needs to be evaluated in further investigations to clarify the economic factor for sustainability in social aspects.

Author Contributions: Conceptualization, Q.H.L. and N.K.; methodology, Q.H.L., H.S. and J.H.; software, Q.H.L., H.S. and J.H.; visualization, Q.H.L.; writing—original draft, Q.H.L.; writing—review and editing, Q.H.L., H.S., N.K. and Y.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 22CTAP-C163903-02) and Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government (MOTIE) (20202020800030, Development of Smart Hybrid Envelope Systems for Zero Energy Buildings through Holistic Performance Test and Evaluation Methods and Fields Verifications).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Rafiq, W.; Musarat, M.A.; Altaf, M.; Napiah, M.; Sutanto, M.H.; Alaloul, W.S.; Javed, M.F.; Mosavi, A. Life Cycle Cost Analysis Comparison of Hot Mix Asphalt and Reclaimed Asphalt Pavement: A Case Study. *Sustainability* **2021**, *13*, 4411. [CrossRef]
2. Pearce, A.; Ahn, Y.H. *Sustainable Buildings and Infrastructure: Paths to the Future*; Routledge: London, UK, 2013; p. 509.
3. Andrade-Núñez, M.J.; Aide, T.M. Built-up expansion between 2001 and 2011 in South America continues well beyond the cities. *Environ. Res. Lett.* **2018**, *13*, 084006. [CrossRef]
4. Chokhachian, A.; Perini, K.; Giulini, S.; Auer, T. Urban performance and density: Generative study on interdependencies of urban form and environmental measures. *Sustain. Cities Soc.* **2020**, *53*, 101952. [CrossRef]
5. Salvati, L.; Gargiulo Morelli, V. Unveiling urban sprawl in the Mediterranean region: Towards a latent urban transformation? *Int. J. Urban Reg. Res.* **2014**, *38*, 1935–1953. [CrossRef]
6. Okazawa, Y.; Murakami, N. *Case Study on Managing Urban Expansion in Tokyo*; World Bank Group: Washington, DC, USA, 2019.

7. Hao, P.; Geertman, S.; Hooimeijer, P.; Sliuzas, R. Spatial analyses of the urban village development process in Shenzhen, China. *Int. J. Urban Reg. Res.* **2013**, *37*, 2177–2197. [CrossRef]
8. Yuan, C.; Chen, L. Mitigating urban heat island effects in high-density cities based on sky view factor and urban morphological understanding: A study of Hong Kong. *Archit. Sci. Rev.* **2011**, *54*, 305–315. [CrossRef]
9. Khamchiangta, D.; Dhakal, S. Physical and non-physical factors driving urban heat island: Case of Bangkok Metropolitan Administration, Thailand. *J. Environ. Manag.* **2019**, *248*, 109285. [CrossRef]
10. Park, B.; Oh, K.; Hong, S. Analysis of the changes in urban thermal environments considering development densities (FAR and BCR). *Int. J. Environ. Sci. Dev.* **2018**, *9*, 32–37. [CrossRef]
11. Usui, H. Variation in building heights under zoning regulations of building coverage ratio and floor area ratio: Theoretical and empirical investigation of downtown districts in Tokyo. In *ISUF 2020 Virtual Conference Proceedings*; University of Utah: Salt Lake City, UT, USA, 2021; Volume 1.
12. Megayanti, T.; Widaningsih, L.; Minggra, R.; Dewi, N. Building Coverage Ratio at the Eastern Corridor of Jalan Ir. H. Djuanda Bandung. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *288*, 012144. [CrossRef]
13. Budiyanto, M.; Prajitno, I.; Hasibuan, H. Sustainable industrial estate by managing the building coverage ratio in Cibitung Industrial Town, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *126*, 012185. [CrossRef]
14. Huang, X.; Cao, Y.; Li, J. An automatic change detection method for monitoring newly constructed building areas using time-series multi-view high-resolution optical satellite images. *Remote Sens. Environ.* **2020**, *244*, 111802. [CrossRef]
15. Nüssli, R.; Schmid, C. Beyond the urban–suburban divide: Urbanization and the production of the urban in Zurich North. *Int. J. Urban Reg. Res.* **2016**, *40*, 679–701. [CrossRef]
16. Soliman, A.; Mackay, A.; Schmidt, A.; Allan, B.; Wang, S. Quantifying the geographic distribution of building coverage across the US for urban sustainability studies. *Comput. Environ. Urban Syst.* **2018**, *71*, 199–208. [CrossRef]
17. Hao, X.; Zhang, G.; Ma, S. Deep Learning. *Int. J. Semant. Comput.* **2016**, *10*, 417–439. [CrossRef]
18. Cao, R.; Tu, W.; Yang, C.; Li, Q.; Liu, J.; Zhu, J.; Zhang, Q.; Li, Q.; Qiu, G. Deep learning-based remote and social sensing data fusion for urban region function recognition. *ISPRS J. Photogramm. Remote Sens.* **2020**, *163*, 82–97. [CrossRef]
19. Maxwell, A.E.; Warner, T.A.; Guillén, L.A. Accuracy assessment in Convolutional Neural Network-based Deep Learning remote sensing studies—Part 2: Recommendations and best practices. *Remote Sens.* **2021**, *13*, 2591. [CrossRef]
20. Maxwell, A.E.; Warner, T.A.; Guillén, L.A. Accuracy assessment in Convolutional Neural Network-based Deep Learning remote sensing studies—Part 1: Literature Review. *Remote Sens.* **2021**, *13*, 2450. [CrossRef]
21. Hu, Q.; Zhen, L.; Mao, Y.; Zhou, X.; Zhou, G. Automated building extraction using satellite remote sensing imagery. *Autom. Constr.* **2021**, *123*, 103509. [CrossRef]
22. Guo, H.; Shi, Q.; Marinoni, A.; Du, B.; Zhang, L. Deep building footprint update network: A semi-supervised method for updating existing building footprint from bi-temporal remote sensing images. *Remote Sens. Environ.* **2021**, *264*, 112589. [CrossRef]
23. Zhao, W.; Persello, C.; Stein, A. Building outline delineation: From aerial images to polygons with an improved end-to-end learning framework. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 119–131. [CrossRef]
24. Meinel, G.; Hecht, R.; Herold, H. Analyzing building stock using topographic maps and GIS. *Build. Res. Inf.* **2009**, *37*, 468–482. [CrossRef]
25. Pan, X.-Z.; Zhao, Q.-G.; Chen, J.; Liang, Y.; Sun, B. Analyzing the variation of building density using high spatial resolution satellite images: The example of Shanghai city. *Sensors* **2008**, *8*, 2541–2550. [CrossRef] [PubMed]
26. Kajimoto, M.; Susaki, J. Urban density estimation from polarimetric SAR images based on a POA correction method. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1418–1429. [CrossRef]
27. Susaki, J.; Kajimoto, M.; Kishimoto, M. Urban density mapping of global megacities from polarimetric SAR images. *Remote Sens. Environ.* **2014**, *155*, 334–348. [CrossRef]
28. Gonzalez-Aguilera, D.; Crespo-Matellan, E.; Hernandez-Lopez, D.; Rodriguez-Gonzalvez, P. Automated rrban analysis based on LiDAR-derived building models. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1844–1851. [CrossRef]
29. Hecht, R.; Kunze, C.; Hahmann, S. Measuring completeness of building footprints in OpenStreetMap over space and time. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 1066–1091. [CrossRef]
30. Emmanuel, B.; Charou, E.; Tsenoglou, T.; Vassilas, N. Automated building block extraction and building density classification using aerial imagery and LiDAR data. *J. Earth Sci. Eng.* **2016**, *6*, 1–9.
31. Koc-San, D.; Selim, S.; Aslan, N.; San, B.T. Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform. *Comput. Electron. Agric.* **2018**, *150*, 289–301. [CrossRef]
32. Feng, Q.; Liu, J.; Gong, J. UAV remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote Sens.* **2015**, *7*, 1074–1094. [CrossRef]
33. Tekouabou, K.; Cedric, S.; Diop, E.B.; Azmi, R.; Jaligot, R.; Chenal, J. Reviewing the application of machine learning methods to model urban form indicators in planning decision support systems: Potential, issues and challenges. *J. King Saud Univ. Comput. Inf. Sci.* **2021**, *34*, 5943–5967.
34. He, S.; Jiang, W. Boundary-Assisted Learning for Building Extraction from Optical Remote Sensing Imagery. *Remote Sens.* **2021**, *13*, 760. [CrossRef]
35. Awan, H.H.; Hussain, A.; Javed, M.F.; Qiu, Y.; Alrowais, R.; Mohamed, A.M.; Fathi, D.; Alzahrani, A.M. Predicting Marshall Flow and Marshall Stability of Asphalt Pavements Using Multi Expression Programming. *Buildings* **2022**, *12*, 314. [CrossRef]

36. Habib, A.; Kim, E.-M.; Kim, C. New Methodologies for True Orthophoto Generation. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 25–36. [CrossRef]
37. Ji, S.; Shen, Y.; Lu, M.; Zhang, Y. Building instance change detection from large-scale aerial images using convolutional neural networks and simulated samples. *Remote Sens.* **2019**, *11*, 1343. [CrossRef]
38. Ma, H.; Liu, Y.; Ren, Y.; Yu, J. Detection of collapsed buildings in post-earthquake remote sensing images based on the improved YOLOv3. *Remote Sens.* **2020**, *12*, 44. [CrossRef]
39. Li, Y.; Hu, W.; Dong, H.; Zhang, X. Building damage detection from post-event aerial imagery using single shot multibox detector. *Appl. Sci.* **2019**, *9*, 1128. [CrossRef]
40. He, K.G.G.; Dollar, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
41. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
42. Qin, X.; Wu, C.; Chang, H.; Lu, H.; Zhang, X. Match Feature U-Net: Dynamic receptive field networks for biomedical image segmentation. *Symmetry* **2020**, *12*, 1230. [CrossRef]
43. Touzani, S.; Granderson, J. Open data and deep semantic segmentation for automated extraction of building footprints. *Remote Sens.* **2021**, *13*, 2578. [CrossRef]
44. Li, J.; Meng, L.; Yang, B.; Tao, C.; Li, L.; Zhang, W. LabelRS: An automated toolbox to make deep learning samples from remote sensing images. *Remote Sens.* **2021**, *13*, 2064. [CrossRef]
45. San-José, J.T.; Garrucho, I.; Losada, R.; Cuadrado, J. A proposal for environmental indicators towards industrial building sustainable assessment. *Int. J. Sustain. Dev. World Ecol.* **2007**, *14*, 160–173. [CrossRef]