Article

Integrating Drone Imagery and AI for Improved Construction Site Management through Building Information Modeling

Wonjun Choi, Seunguk Na and Seokjae Heo *

School of Architecture, Dankook University, Yongin-si 16890, Gyeonggi-do, Republic of Korea;
martin_choi@dankook.ac.kr (W.C.); drseunguk@dankook.ac.kr (S.N.)
* Correspondence: mill@dankook.ac.kr

Abstract: In the rapidly advancing field of construction, digital site management and Building Information Modeling (BIM) are pivotal. This study explores the integration of drone imagery into the digital construction site management process, aiming to create BIM models with enhanced object recognition capabilities. Initially, the research sought to achieve photorealistic rendering of point cloud models (PCMs) using blur/sharpen filters and generative adversarial network (GAN) models. However, these techniques did not fully meet the desired outcomes for photorealistic rendering. The research then shifted to investigating additional methods, such as fine-tuning object recognition algorithms with real-world datasets, to improve object recognition accuracy. The study's findings present a nuanced understanding of the limitations and potential pathways for achieving photorealistic rendering in PCM, underscoring the complexity of the task and laying the groundwork for future innovations in this area. Although the study faced challenges in attaining the original goal of photorealistic rendering for object detection, it contributes valuable insights that may inform future research and technological development in digital construction site management.

Keywords: drone; PCM; construction site; digitalization; object detection

1. Introduction

Construction site management comprehensively covers tasks from the initial design phase to construction and maintenance, and the need for digital construction site management is increasingly growing. In the early stages of the construction industry, paper blueprints were used, but with the advent of computer-based software like CAD, the storage and processing of architectural information have become digitalized. This made it possible to overcome the limitations of 2D blueprints and allowed for the intuitive understanding and management of building structures using 3D models.

Since the early 2000s, Building Information Modeling (BIM) technology has been developed in response to the needs of the construction industry. It has evolved from a tool that merely aids visual elements to an advanced tool capable of performing optimization and estimation tasks [1]. Although BIM has brought about innovative changes in the construction industry, its application is limited in construction, site management, and maintenance fields due to its high dependency on on-site data [1–3].

Considering that the initial design phase accounts for 0.4% of the life cycle cost of a building, the construction phase accounts for 16%, and the maintenance phase accounts for 60–80%, the introduction of BIM technology to site management is essential for the construction industry to successfully respond to the fourth industrial revolution [4–6]. However, the implementation is being delayed due to the lack of technology for creating digital models that adequately reflect on-site information.

Digitalization research at construction sites typically involves the use of point cloud model (PCM) methods using robots and scanning equipment [7–10]. PCM represents objects or spaces using a 3D point dataset in space [11]. However, robots and scanners

are expensive equipment and are not suitable for real-time site management due to their susceptibility to on-site conditions [12–14]. Furthermore, to recognize objects in the point cloud form for on-site safety management, as targeted in this paper, an additional process of training a new artificial intelligence model is required. The development of a construction site-specific artificial intelligence model for photorealistic rendering, which has not yet been developed, is essential. Nevertheless, in the near future, construction site management using artificial intelligence will play a crucial role in supporting and optimizing major stages of building construction [15]. Digital construction site management technology, in particular, will act as a key factor in improving efficiency and accuracy in these processes [16]. In addition, construction site management through digitalization technology will provide new possibilities for direct financial benefits to construction companies, thereby contributing to the growth of the construction industry [17,18]. Therefore, construction site digitalization technology must be easily accessible to small and medium-sized enterprises. To this end, this study proposes a method to digitalize construction site information using only drones, excluding expensive sensors.

Drones are used to take photos and videos at construction sites, allowing real-time monitoring of construction project progress. In this study, drone photos will be used to construct PCM models and conduct object recognition research through model photorealistic rendering. This will efficiently establish the digitalization process at construction sites and provide technology that is easily accessible to small and medium-sized enterprises. In addition, considering the short flight time of drones, this study will compare and analyze data collection methods for PCM model construction in a short period. This will reduce costs incurred during construction and maintenance stages at construction sites and improve the overall efficiency of construction projects. Moreover, this study aims to develop technology that connects objects to be recognized through image preprocessing techniques such as model photorealistic rendering and completely different forms of training data that can be used together [19]. Digital technology utilizing artificial intelligence can enhance on-site management's efficiency and accuracy and contribute to innovation and growth across the construction industry. The results of this study are expected to significantly help improve construction site safety and construction quality and reduce project schedules and costs.

In conclusion, this study proposes an efficient and economical construction site management solution for small and medium-sized construction companies by combining drone and digitalization technologies. As future research directions, it is necessary to develop a more precise data collection and analysis system by integrating additional sensors and technologies into drones. This will further enhance the monitoring and management capabilities at construction sites. In addition, the development of such technologies will lay the foundation for achieving sustainable growth in the overall construction industry.

Research Questions and Objectives

This study enhances digital construction site management by harnessing the synergy between drone imagery and artificial intelligence (AI) within the context of Building Information Modeling (BIM). The foundation of this research is laid upon a series of research questions and objectives meticulously crafted to dissect and understand the multifaceted role of drone technology in revolutionizing the construction management landscape, particularly through the lens of point cloud models (PCMs) and object recognition capabilities.

Research Questions:

The first question probes the integration of drone imagery into the digital construction site management process for creating PCM models with enhanced object recognition capabilities. This inquiry aims to unravel the practical applications of drone technology in capturing real-time data and their seamless integration into BIM for bolstering construction management practices.

Subsequently, the research delves into the challenges and limitations inherent in current photorealistic rendering techniques within the domain of PCM. By scrutinizing methodologies such as blur/sharpen filters and generative adversarial network (GAN)

models, this question seeks to identify viable pathways to surmount these obstacles and amplify object detection accuracy in construction site management.

Last, the investigation extends to the ramifications of fine-tuning object recognition algorithms with real-world datasets. This line of inquiry is pivotal in evaluating how the precision and reliability of object detection within construction sites are influenced by the application of real-world data for algorithm optimization.

Research Objectives:

Aligned with the research questions, the objectives of this research are structured to provide a comprehensive examination of the potential and challenges associated with integrating drone imagery and AI in construction site management. The initial objective focuses on evaluating this integration for the creation of PCM models that boast superior object recognition capabilities. Through this evaluation, the study aims to enhance practicality, efficiency, and overall improvements that drone-captured imagery could usher into BIM creation and construction site management.

In pursuit of addressing the second question, the research aims to identify and mitigate the limitations facing current photorealistic rendering techniques in PCM. This endeavor involves a thorough investigation into the prevailing challenges and the proposition of innovative solutions designed to elevate object detection accuracy.

Finally, the objective underscores the enhancement of object recognition algorithms by fine-tuning them with real-world datasets. This approach is anticipated to refine the accuracy and reliability of AI-based object recognition algorithms, leveraging real-world construction site data as a pivotal resource for algorithmic improvement.

Through these research questions and objectives, this study aspires to forge a path toward innovative advancements in digital construction site management, setting a new paradigm in how construction projects are planned, executed, and monitored in the era of drone technology and artificial intelligence.

2. Literature Review

2.1. Construction Site Management

Research for digital construction site management has been conducted by many scholars. As a result, enabling technologies related to BIM, sensors, unmanned aerial vehicles (UAVs), and unmanned ground vehicles (UGVs) have been developed for use on-site [20,21]; however, there are still no widely adopted commercial technologies in the construction industry. The biggest issue in realizing digital site management is the absence of a digital twin (DT) for the construction site [22–25]. Due to such problems, the enabling technologies prepare only the minimum necessary information in different formats, making it difficult for them to share information organically. For successful digital construction site management, there is a need for a single integrated model and platform that enables collaboration between these enabling technologies. In this section, we aim to examine the current state of research in construction management, safety management, and maintenance, which constitute a major part of construction site management, to more clearly understand why the DT has not yet been realized and to propose solutions.

During the construction phase, research and development focus on process efficiency, production efficiency, and monitoring logistics and work progress [23,26]. For process efficiency, the trend is to concentrate on minimizing defects caused by various factors and delays resulting from them. A representative method is to precisely scan site information using equipment and sensors and prefabricate optimized frames and components through off-site construction (OSC). These components are later assembled on-site. Since they were produced in a factory setting with fewer external influences, defects are reduced, rework frequency decreases, and ultimately, construction time is shortened [27,28]. Furthermore, reducing defective components implies a proportional reduction in the time needed for reproduction, equivalent to increased production efficiency. Additionally, AI-powered logistics and work progress monitoring technologies have proven effective through pilot projects [5]. Site data collection primarily utilizes UAVs and UGVs equipped with sensors

and the collected data are later reconstructed into a PCM. After converting site data into PCM, object recognition is conducted using algorithms or AI, and the resulting digital twin model is highly accurate. However, as mentioned in the introduction of this paper, the current process is not suitable for monitoring constantly changing sites due to the time and effort required to prepare the data necessary for AI training [29,30], necessitating a different solution.

As societal concern for safety increases, there is a growing demand for technology and research related to safety management. Current safety management involves a safety officer assessing risk factors for each task, taking appropriate measures, and monitoring whether the site adheres to safety regulations [31]. However, as the scale of construction projects increases, so does the number of tasks, and the number of safety officers is insufficient [32]. Therefore, research in safety management mainly aims to automate the safety management process, minimize blind spots, and enable a small number of safety officers to oversee a larger site. Representative technologies include AI and algorithm-based risk factor identification, real-time object tracking, and worker behavior pattern analysis [33–35], all of which involve recognizing objects, including people and equipment. However, existing research lacks practicality when considering efficiency [36,37]. For example, research enabling highly accurate site monitoring and safety management requires many expensive sensors, while technologies that can cover the entire construction site are not accurate enough. Based on construction accident statistics, the most common accidents vary by country but include falls, entrapments, collisions, crushes, and being struck by falling objects [38,39], all of which are difficult to predict and happen in an instant. For successful safety management, technology that can quickly process and monitor real-time data and accurately identify worker behavior is necessary.

In terms of maintenance-related research, it can be divided into two main directions: improving operational efficiency and ensuring stability through preventive maintenance. First, looking at research on operational efficiency improvement, the goal is to reduce costs as most of the life cycle costs of a building occur during the period of use, from completion to demolition. As operational efficiency can be optimized by clearly understanding information about the building, attempts are actively made to detect and analyze factors affecting operation, such as temperature, humidity, and circulation, using IoT sensors [22,40–42]. Recently, there has been a trend of introducing artificial intelligence to discover correlations between data that are difficult for humans to intuitively grasp and to predict future data based on past data [43]. In addition, maintaining the functionality of a building is also important in maintenance, so preventive maintenance technologies developed through real-time monitoring are being developed. However, while research on operational efficiency improvement through data analysis is approaching commercialization, preventive maintenance technology is relatively lacking. The main reason is that monitoring technology is still in the research stage [44], and new technology development that can be used for internal and external maintenance of buildings is needed.

Last, in all three areas of construction management, maintenance, and safety management, various attempts are being made using sensors and unmanned vehicles (UVs) to digitize on-site information. Sensors show significant differences in price and performance depending on the type, and as they cannot move without human intervention, data collection methods using unmanned vehicles (UVs) are gaining attention [20]. In particular, among UVs, UAVs have a short flight time due to technical limitations, so UGVs are mainly used on-site [12,14,40,45,46]. However, UGVs used to collect on-site data are often expensive, although there are differences depending on the type. For example, Amazon's Spot robot, which can move around a construction site automatically while overcoming obstacles, costs about $75,000 per unit. From a cost perspective, this can be a factor that hinders the adoption of technology as it is unclear how much cost-saving effects can be obtained from digitizing on-site information and how many UGVs need to be used [47]. Therefore, it is considered realistic to use relatively low-cost UAVs for data collection in

terms of technology acceptance. However, a process capable of collecting sufficient data in a short period of time is required to overcome the inherent short flight time.

2.2. Advancements in PCM to BIM Conversion, Photorealistic Rendering, and AI-Driven Object Recognition

The conversion from point cloud models (PCMs) to Building Information Models (BIM) is essential in digital construction site management. Existing methodologies utilize automated algorithms for object identification within PCMs, followed by semantic enrichment and organization into a structured BIM model [48,49]. However, these methodologies often require significant manual intervention, especially in the semantic enrichment and data synchronization stages, which can lead to potential inaccuracies and inefficiencies in the generated BIM models [50]. The manual processes also extend the time required for PCM to BIM conversion, which is not suitable for dynamic construction environments where timely updates are crucial [51].

In photorealistic rendering, existing techniques like ray tracing and radiosity have contributed to enhanced visual quality but may not achieve the level of realism necessary for stakeholders in construction projects [52]. The visual appeal and informational value of models generated through these techniques may lack the required detail and accuracy for applications like virtual site inspections and client presentations [53]. Additionally, the generation of realistic textures remains a challenge not adequately addressed in the existing literature [54].

Regarding the advancement of artificial intelligence (AI) in object recognition within vision processing, significant progress has been made [55]. However, the application of AI, especially in complex construction site environments, often requires substantial computational resources and meticulously curated datasets for training the models [56,57]. The challenge of collecting and preparing high-quality training data, coupled with the demand for high computational power, is a significant barrier to the widespread adoption of AI-driven object recognition in construction site management [58].

3. Research Method

3.1. Process

In the introduction, as previously presented, a method was proposed to perform object recognition by creating a PCM model using drone technology and then creating a photorealistic rendering of the model. In this section, the newly developed process is compared with the existing process, and the differences between the two processes are confirmed.

As shown in Figure 1, the prevailing site management methodology leverages robots and sensors to amass scanning data from the site. This collected data serve as a bedrock for constructing a Point cloud model (PCM), subsequently necessitating the preparation of training data for object recognition [57,59]. The process of object recognition is orchestrated through a dedicated artificial intelligence training regimen. Nonetheless, this procedure presents significant hurdles, chiefly the labor-intensive task of preparing training data in PCM format—a format not easily accessible. The extensive duration extending from scanning to object recognition significantly detracts from its applicability in real-time site management, marking a notable limitation of this conventional approach.
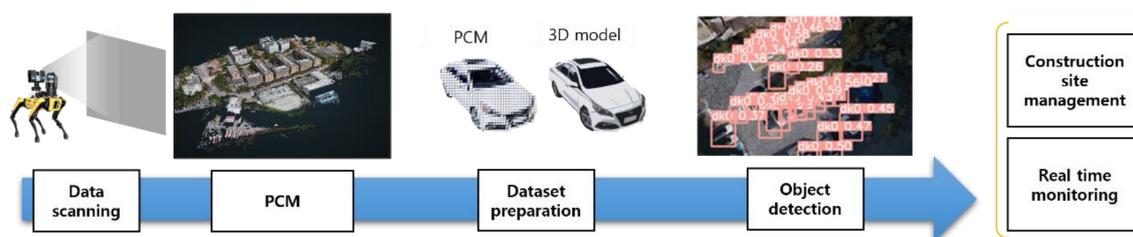

Figure 1. The original process of construction site digitalization.

The method proposed in this study uses drones instead of existing robots and sensor technology to collect data, and the new process described in Figure 2 aims to reduce the time and cost required to build a training dataset. The improved site management process collects data through drone photography and constructs a PCM using the WebODM library with the collected photos. WebODM was selected due to its compatibility with various drone technologies, allowing for the efficient processing of aerial photographs into a 3D model. Its ability to handle large datasets while maintaining precision is well suited for the construction of detailed PCM models from drone imagery. Furthermore, WebODM provides a user-friendly interface, enabling fine-tuned customization over processing parameters. This control allows the PCM construction to align with the specific requirements of the project. Then, the whole or necessary partial areas for object recognition are extracted from the PCM model, and the extracted images are converted to photorealistic images using generative adversarial network (GAN) AI.
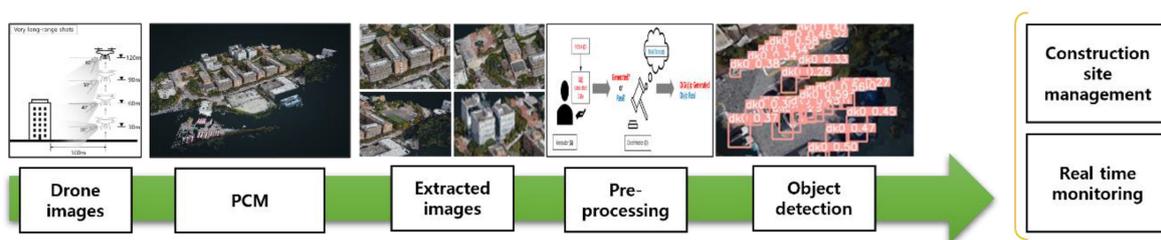
Figure 2. The proposed process of construction site digitalization.

The main shift from the existing process after creating the PCM is that object recognition is performed after generating a photorealistic rendering of the PCM model with GAN. This step aims to reduce the computational resources needed for site management, which is crucial for successful construction site management. When a point cloud-based model is crafted at a construction site filled with many irregular objects for artificial intelligence use, the most time-consuming task is setting up the training dataset to detect objects from the PCM. Moreover, even after the artificial intelligence capable of recognizing point clouds is ready, additional fine-tuning of the training dataset in PCM format is needed for each site due to different shapes, which lowers its practicality [60–62]. However, if the point cloud model is converted into a photorealistic rendering as proposed in this study, identifying equipment and materials at the site using commonly used object recognition artificial intelligence becomes much easier. Furthermore, this approach can significantly speed up the process of creating BIM models through improved object detection from PCMs.

To achieve a photorealistic rendering of the PCM model, a variant of the generative adversarial network (GAN), a generative artificial intelligence model, was employed, inspired by the concept of transforming sketches into images. This selection emerged from a meticulous evaluation process. Other contemporary models such as Stable Diffusion models and Variational AutoEncoders (VAE) were also contemplated but were sidelined due to their inconsistencies in object transformation during the generation phase. This limitation underscored the need for a model that embodies both consistency and reliability. The evaluation then pivoted toward analyzing the quality of generation where GANs surfaced as the most fitting candidate. GAN surpassed other models by displaying superior performance in aspects like generation quality, training speed, and stability [63]. Its prowess in crafting lifelike images deemed it apt for photorealistic rendering, aligning seamlessly with the project's objectives.

Expanding on this aptitude, the versatility of GAN enabled it to address specific requirements within the PCM environment, ensuring a faithful representation of intricate details. The capability to generate high-resolution and realistic images was a pivotal factor in the decision-making process, with GAN's performance in these domains distinguishing it from other models.

The adoption of the GAN model was a judicious decision grounded on a thorough assessment of its performance, flexibility, and congruence with project goals. It was chosen not merely for its ability to transmute sketches into images but also for its demonstrated proficiency in generating realistic images and its adaptability to diverse scenarios. Hence, GAN emerged as the most suitable choice for generating photorealistic rendering from sketches.

Furthermore, with an eye toward real-time site management, the expeditious and precise YOLOv5 was selected for object recognition from a range of vision-processing artificial intelligence tools. YOLO, an acronym for You Only Look Once, is a real-time object detection artificial intelligence framework that predicts bounding boxes and class probabilities of objects in images utilizing a singular convolutional neural network (CNN). The notable speed and high accuracy of YOLO render it apt for construction site management, which often entails dealing with a diverse array and quantity of objects [64]. Since its inaugural release in 2015, the YOLO model has undergone iterative enhancements, leading to the subsequent versions YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, YOLOv7, and YOLOv8, with the attributes of each version encapsulated in Table 1.

Table 1. Explanation of YOLOv1, v2, v3, v4, v5, v6, v7, and v8.

| Version		Characteristics		Note
YOLOv1	-	Utilizes a single-convolution neural network (CNN) to represent object class probabilities and determine object regions as bounding boxes	- -	High processing speed Lower accuracy compared with subsequent versions
YOLOv2	-	Employs the Darknet-19 convolutional neural network and a "logistic regression loss" loss function	- -	Higher accuracy compared with YOLOv1 Capable of recognizing more object classes
YOLOv3	- -	Uses the Darknet-53 convolutional neural network and a "focal loss" loss function Focal loss focuses on learning hard-to-recognize objects to improve accuracy	- -	Higher accuracy compared with YOLOv2 Can detect smaller objects
YOLOv4	- -	Implements the CSPDarknet-53 convolutional neural network and a CIoULoss loss function, which offers higher accuracy than IoU Applies mosaic data augmentation	- - -	Higher accuracy compared with YOLOv3 Can detect even smaller objects Faster processing speed
YOLOv5	- -	Utilizes anchor boxes with varying ratios to enhance accuracy Applies random cropping and flipping augmentation techniques	- -	Lower capacity compared with YOLOv4 Easily recognizes small objects in wider areas
YOLOv6~8	- - -	Suggests trainable bag-of-freebies Used quantization and distillation methods to improve accuracy A new repository for YOLO models has been launched for training object detection, instance segmentation, and image classification.	- -	Higher accuracy compared with YOLOv5 Larger model compared with YOLOv5

Construction site management encompasses a broad spectrum, necessitating a model proficient in wide-area recognition. Furthermore, given its intrinsic link to safety, a high degree of accuracy in object recognition is indispensable, and the leeway for delays engendered by sluggish processing speeds is scant. In light of these requisites, this study opted for YOLOv5, revered for its expansive area scanning capability, superior object recognition accuracy, and brisk processing speeds among the YOLO iterations. Although the later versions of YOLO, specifically YOLOv6, 7, and 8, manifest higher accuracy compared with YOLOv5, their larger model sizes levy a higher demand on computational resources [64,65]. Consequently, YOLOv5 represents an optimal balance between model size and accuracy, aligning well with the project's objectives.

In summary, this research aimed to address the issues arising from the conventional methods by applying a drone-based data collection process and object recognition through photorealistic rendering. Additionally, the study proposes a new process to enhance the safety and efficiency of construction sites by utilizing fast and accurate artificial intelligence algorithms, such as YOLOv5, for real-time site management.

3.2. PCM Generation Using Drones

First, in the existing process, high-priced Lidar sensors were combined with drones to create PCM models. This method enables the creation of precise PCM models by approaching the building as closely as possible due to the measurement limitations of the Lidar sensors. However, it is impossible to recognize the entire building at once due to the need for proximity. Although the technology has reached the point of practicality, it has become unusable due to equipment issues and the short flight time of drones.

Digital construction site management requires rapid, periodic site model generation [66,67]. Therefore, it was determined that the technology to create a site model based on images collected using drone advantages such as fast movement, wide-angle photography, and recording of shooting locations is practical. In particular, if swarm drones are used to assign specific areas of the site, as shown in Figure 3, the efficiency will increase proportionally. Therefore, research on technology to ensure the accuracy of PCM while securing a minimum amount of data in the image collection process was needed for optimal process construction.
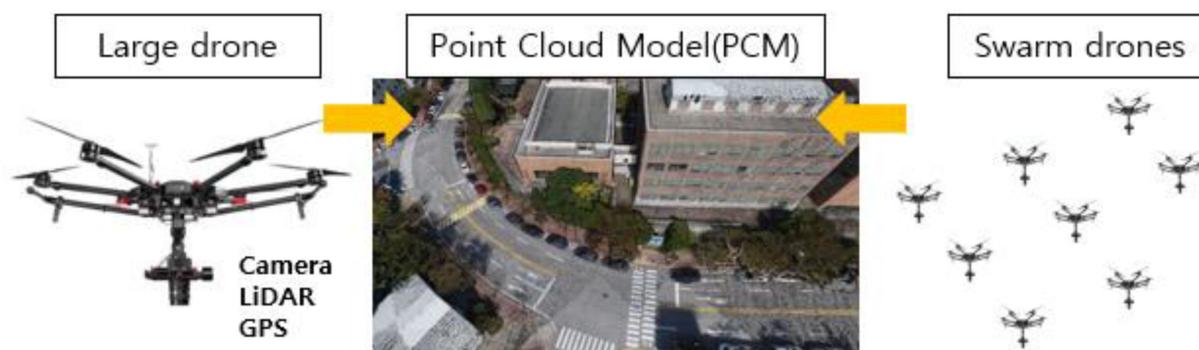

Figure 3. Explanation of the technology that creates PCM using only images by collecting images while automatically moving to designated locations using swarm drones on the right, compared with the traditional method of creating a PCM model in the center by manually controlling a large drone with various sensors on the left.

In this section, a case study was conducted to determine the optimal conditions for a drone to acquire site photos by adjusting altitude, angle, and distance while considering the overlap rate, as shown in Table 2. The overlap rate is crucial for creating 3D models and refers to the proportion of overlap between the collected photos, as illustrated in Figure 4. Insufficient overlap between photos can result in errors during PCM generation, leading to distortions. Both longitudinal and lateral overlaps are required.

The drone utilized for image acquisition in this study was DJI's Mavic Mini 2, a lightweight model weighing 249 g and priced at around USD 500–600, with a flight duration of approximately 20 min. It was selected due to its adjustable camera shooting angle although its GPS-based automatic movement feature was not utilized in this study. The image data collected were categorized into vertical, close-, long-, and very long-range shots. PCM models crafted based on the captured images were compared and analyzed. Importantly, the Mavic Mini 2 is capable of capturing high-resolution images with a resolution of 12 megapixels and videos at 4 K/30 fps. This high resolution is integral for accurate PCM generation. Despite its cost-effectiveness, the image resolution offered by the Mavic Mini 2 proved to be adequately suitable for the object detection tasks required in this study, illustrating a favorable balance between cost and quality in the chosen UAV model.

Table 2. Summary of cases according to image collection methods.

Case	Shooting Distance (m)	Shooting Altitude (m)	Shooting Angle from Drone (°)	Number of Images	Time (Sec)	Resolution	Visual Judgment Result
Vertical shooting	0~20	70	90	110	420	2.5 cm/px	Building well recognized; windows on the side and the walls are not well modeled
Close-range shooting	10	15~60	30	110	510	0.4 cm/px	Building not recognized
Close-range and angled shooting	10	15~60	15~60	108	630	~1.6 cm/px	Building not recognized
Long-range shooting	60	60	45	102	300	2.2 cm/px	Building well recognized
Long-range and angled shooting	30	20~60	20~60	120	580	~2.5 cm/px	Building well recognized with less empty space
Very long-range shooting	100	30~120	30~60	196	460	~5 cm/px	Building well recognized except for hard-to-identify small objects
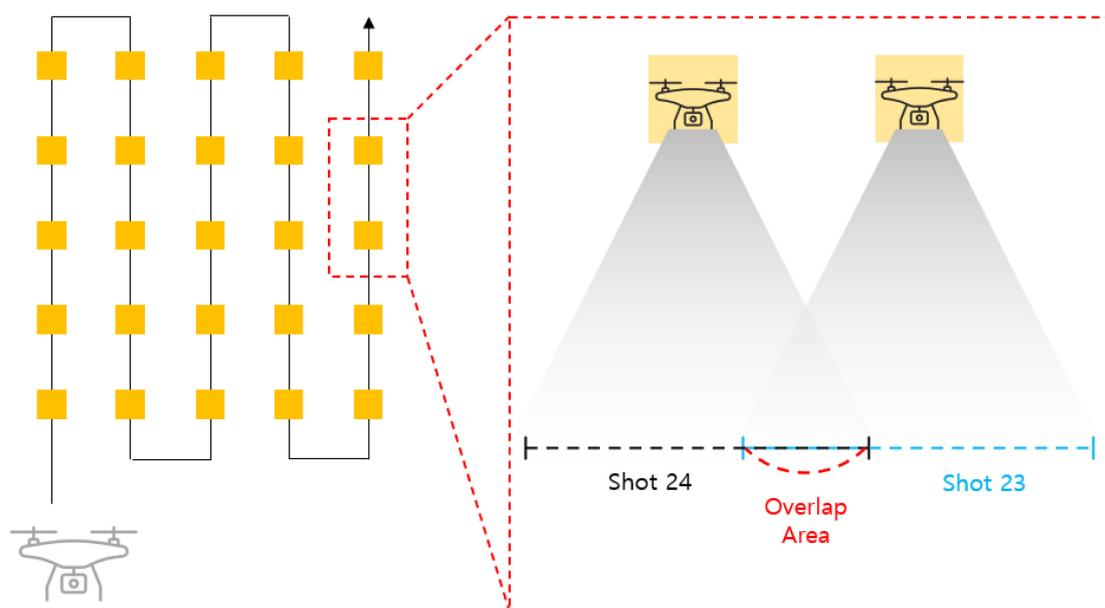
Figure 4. Diagram of overlap area between drone-captured images for 3D model generation. The image on the left shows the drone's movement path, and the orange squares represent the positions where photos are taken along the path. The right image demonstrates what the "overlap" means when capturing photos.

The drone's altitude and angle were calculated based on the explained overlap rate. The overlap rate criteria are based on the "Drone Cadastral Survey Manual", shared during a conversation with officials from the "Korea Land and Geospatial Informatix Corporation," as shown in Table 3. The collected images were taken by the same researcher without any post-processing and were collected as JPG files. The shooting was performed by rotating around the target building and repeating the process, which is the same method as automatically capturing images while descending in a specific area to match the future on-site BIM creation automation process.

Table 3. Criteria for overlap rate in drone cadastral surveys based on the Korea Land and Geospatial Informatix Corporation manual.

Category	Flat Area	Area with Elevation Difference	Insufficient Matching Points	Area with High-Rise Buildings
Longitudinal overlap rate	65%	75%	75%	85%
Latitudinal overlap rate	60%	70%	70%	80%

Shadows occurred on the building depending on the shooting time, but that part was not considered and was used as a dataset. The reason is to determine whether the recognition is possible even if a slight shadow hides the building or objects, and ultimately, the shadow did not affect the research progress as a variable. Since the goal of this study is to check whether objects are recognized within the created model, not the precise model creation for safety diagnosis or inspection, the resolution was not analyzed. The PCM created by LiDAR sensors also cannot collect sensitive building information such as mold or wall discoloration unless it is very close, as the resolution is 3 cm to 10 cm when collecting data from a long distance. This point was considered during the analysis.

The shooting target was set as the university building where the researcher is affiliated, the reason being that it was difficult to obtain permission for drone shooting for research purposes at a construction site due to safety issues. Instead, a site similar to an apartment construction site was chosen where a building exists in a particular shape and there is a parking lot nearby where vehicles pass. Therefore, the research target was the buildings around the College of Engineering at Dankook University. The research building consists of four 5-story buildings in a row, with roads surrounding the building and vehicles parked on the road, making it a suitable target for the AI to recognize buildings, roads, vehicles, and other objects. For medium- and long-range shooting, the entire school premises were captured, as shown in Figure 5, and a PCM was created, including the target buildings.

Figure 5. Exemplary captured images, shooting while rotating around the building.

3.3. Vertical Shooting Dataset

We assessed whether the entire building could be converted to a PCM by taking vertical shots of the target building. The purpose was to determine if images taken vertically, similar to those from satellites, could be used for research. Image collection was conducted at an altitude of 60 m, and as shown in Figure 6, the drone moved horizontally at regular intervals while shooting. As a result, the overall shape appeared well in Figure 7, but there was a lack of information on vertical surfaces, and when zooming in on the model, the expression of the building's side was inadequate. Nevertheless, the overall shape of the building was well represented, and it was judged that vertical shooting data could be used under limited circumstances.
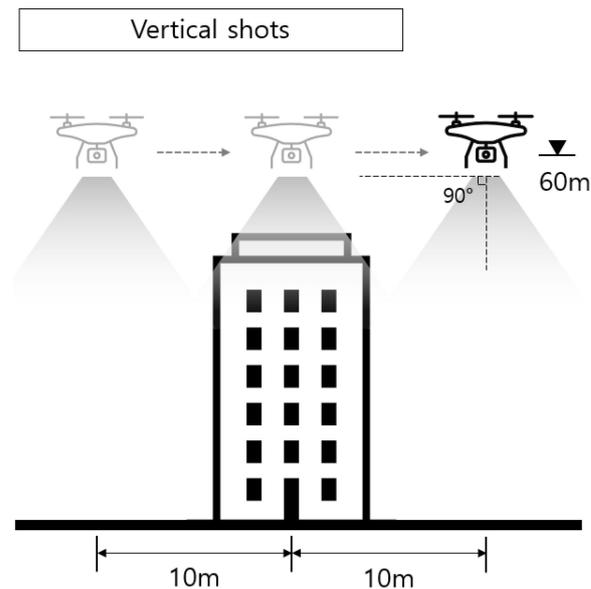
Figure 6. Vertical shots.

Figure 7. Vertical shot PCM data.

3.4. Close-Range Shooting Dataset

We examined two cases of close-range shooting. First, the building was shot from a horizontal distance of 10 m, raising the altitude by 15 m at a time, as shown in Figure 8. Second, to check the impact of the camera lens angle on the model's completion, the lens angle was changed by 15 degrees at the same distance as Figure 9. Ultimately, the building modeling failed, as shown in Figure 10. This failure was judged to be due to a lack of correlation information about the target building in each image, making it difficult to clearly understand the relationship between the images.
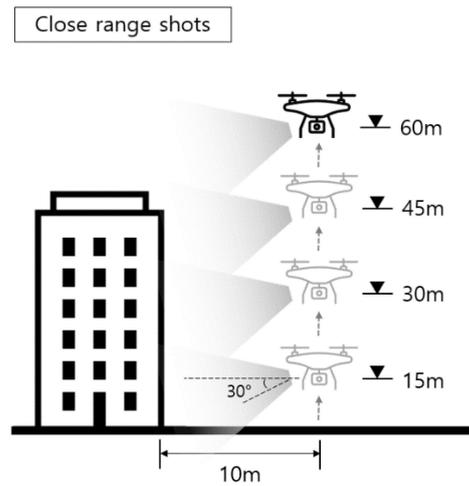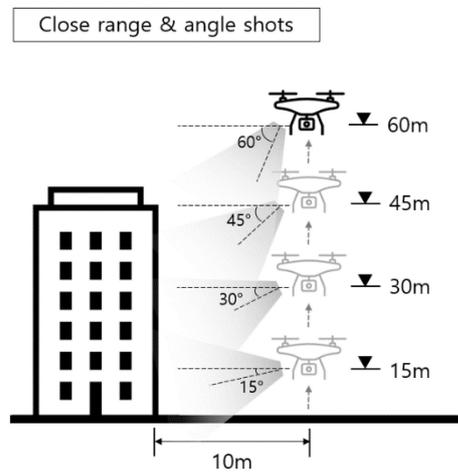
Figure 8. Close-range shots.

Figure 9. Close-range and angle shots.

Figure 10. Close-range shot **PCM** data.

3.5. Long-Range Shooting Dataset

In long-range shooting, we used a method of shooting at regular altitudes and simultaneously changing altitudes and lens angles. Long-range shooting images have all the objects to be recognized within the camera's field of view, so the number of altitudes and lens angle changes is relatively low, as shown in Figures 11 and 12. **PCM** models created based on long-range shooting images showed higher overall completeness than those

created with close-range shooting images. However, the completeness of small objects was relatively lower, as shown in Figures 13 and 14.
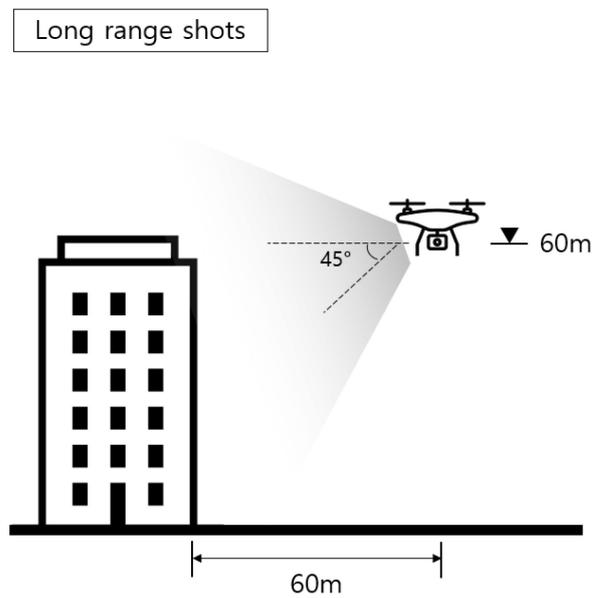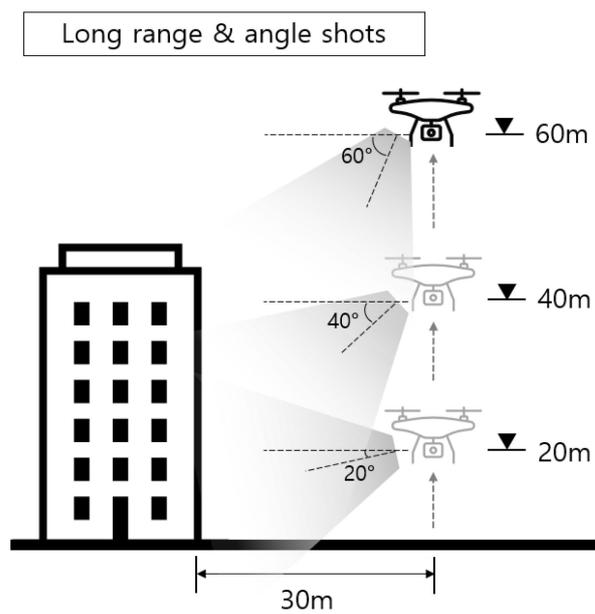

Figure 11. Long-range shots.
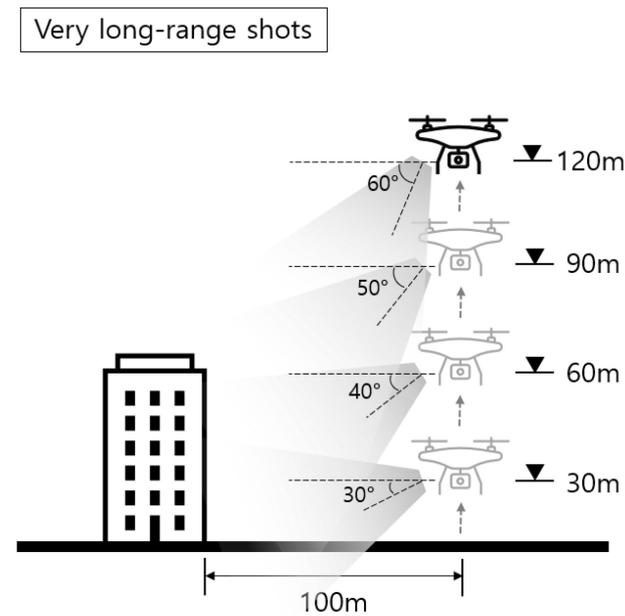

Figure 12. Long-range and angle shots.

Figure 13. Long-range shot **PCM** data.

Figure 14. Long-range and angle shot **PCM** data.

3.6. Very Long-Range Shooting Dataset

The very long-range shooting was conducted at various altitudes based on a horizontal distance of 100 m, as shown in Figure 15. In very long-range shots, each image contained a wide range of fields, but small objects other than buildings appeared challenging to distinguish. Various altitudes and lens angles were used as the basis for image collection, and as a result, noise from small objects increased in the **PCM** model, as shown in Figure 16. Considering that the accuracy of the **PCM** model based on long-range shooting images was higher than in other cases, it was confirmed that shooting should occur at an appropriate distance.

Figure 15. Very long-range and angle shot PCM data.

Figure 16. Very long-range and angle shot **PCM** data.

4. Discussion

In this chapter, we investigate methods for the photorealistic rendering of the generated PCM model. We considered two techniques for the photorealistic rendering of the model: the super resolution technique and the generative adversarial network (GAN). Super resolution technique is a technology that transforms low-resolution images into high-resolution images and is suitable for enhancing the quality of images when textures have already been applied to the model. However, the super resolution technique is used to enhance the pixel resolution and does not inherently add realistic textures or other photorealistic qualities to the images. Therefore, for this study, which uses PCM technology that consists of a simple collection of 3D points, it was determined that using GAN to generate images similar to actual building photos and applying real building textures to the PCM model would be more advantageous. Therefore, this chapter summarizes the research results performed using GANs for photorealistic rendering of buildings.

4.1. Data Preprocessing and Noise Reduction

Before the actual model rendering, object recognition was performed on the non-rendered model to create a control group. As a result, as shown on the left of Figure 17, it was confirmed that AI, which was trained only with general images, could not recognize objects in PCM without rendering. The cause was determined to be the low recognition rate due to empty spaces between points in the on-site point cloud model acting as noise. Therefore, a preprocessing filter was applied to reduce noise and solve the problem of unrecognized objects. Furthermore, although there are various objects at the actual construction site, only buildings and cars were selected to be recognized in this research stage since the purpose is to identify the possibilities of new processes.


Figure 17. YOLO object detection from PCM without filter (**left**); YOLO object detection from PCM with filter (**right**).

To fill the empty spaces between points in PCM, it is necessary to expand the area of all points. For this, blur and sharpen filters were used concurrently for model rendering. As shown in Figure 18, the blur was applied first to expand the area of each point within the point cloud, and the sharpen function was used to correct the blurred image to make it clearer. The most widely used blur types—average, median, bilateral, and Gaussian—were used. Average blur divides the image into square areas and blurs it by calculating the average value with surrounding areas. Median blur removes noise during the blurring process. Bilateral blur considers the color and distance of the image, maintaining sharp areas while blurring, and Gaussian blur uses the Gaussian function to blur the image.

Furthermore, in cases where the gaps between points are wide, a single filter may not be enough to sufficiently fill the surrounding empty space so that AI can recognize it. Therefore, filters were applied a minimum of 1 to a maximum of 6 times, and the sharpen function was always used before applying the next blur to prevent excessively blurred situations where objects could not be recognized. Moreover, by varying the type and order

of blur when repeatedly applying the blur, the optimal blur/sharpen filter combination for model rendering was found.

The mean squared error (MSE) is one of the widely used comparison metrics in image processing, which calculates the average of the squared differences between pixel values of the original image and the filtered image. Therefore, a smaller value means the two images are more similar. The peak signal-to-noise ratio (PSNR) is a variation of the MSE, representing the signal-to-noise ratio between the original and filtered images. PSNR is usually expressed in dB units; a higher value means better image quality. These techniques were used to compare filtered data visually, allowing for structural similarity to be compared as well. To calculate these metrics, the original image and the filtered image were directly compared, and the scikit-image library in Python was used for calculations.
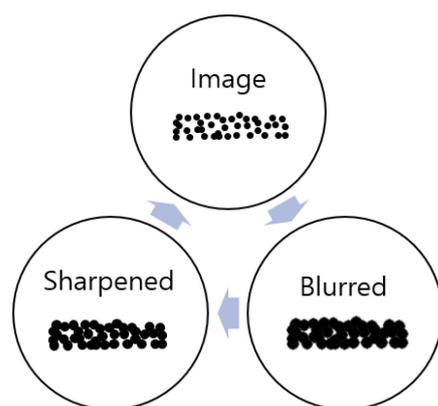
**Figure 18.** Blur/sharpen filter process.

As a result of applying the blur/sharpen filter, as partially shown in Figure 19b–f, the degree of photorealism in the rendered model varied significantly depending on the combination and the number of repetitions of blur types. The MSE and PSNR values in the images are the average values of several images. When comparing filter performance, the similarity order is determined to be c-d-b-e-f. This result also affected the YOLO AI object recognition results, as shown in Figure 17, where applying the filter increased the recognition rate from 0% to 66%. However, the photorealistic rendering created using the blur/sharpen filter tended not to recognize smaller objects, and the increase in recognition rate occurred mainly for relatively larger objects. Moreover, not all combinations of blur filters resulted in an increased recognition rate, and there were cases where the recognition rate decreased depending on the filter combination.

As mentioned in the introduction, the YOLO AI used for recognition does not require a separate data refinement process and utilizes quickly accessible images from the internet. The training dataset consists of ordinary photos from the "PKLot Image Dataset" [68], and "facadesArchitecture Image Dataset" [69], as shown in Figure 20, and can be easily expanded through data crawling, depending on the situation.

Based on the results, further research was conducted on the method to create images most similar to actual photos using filters alone. The four blur filters—average, median, bilateral, and Gaussian—were applied with repetitions ranging from 1 to 6 times. The number of combinations increased by 4^6 for each iteration, resulting in 4096 combinations. Object recognition was performed for all combination cases, and the average recognition accuracy values were calculated and summarized in the scatter plot shown in Figure 21. Each plot in the scatter plot represents the average accuracy obtained from object recognition using a specific combination. The results indicate that excessive filter combinations significantly reduce the correlation between images, and there seems to be a limit to achieving photorealistic rendering using noise filters alone.

Even so, it was confirmed that the possibility of better object recognition with a photorealistic rendering of PCM models created through filter application, without additional

fine-tuning to the AI, is higher than with the original PCM models. However, the 45–61% object recognition accuracy is insufficient for practical use in on-site management. This result seems to be due to the low resolution of the PCM models, and a PCM model closer to actual photos is needed for successful on-site management.
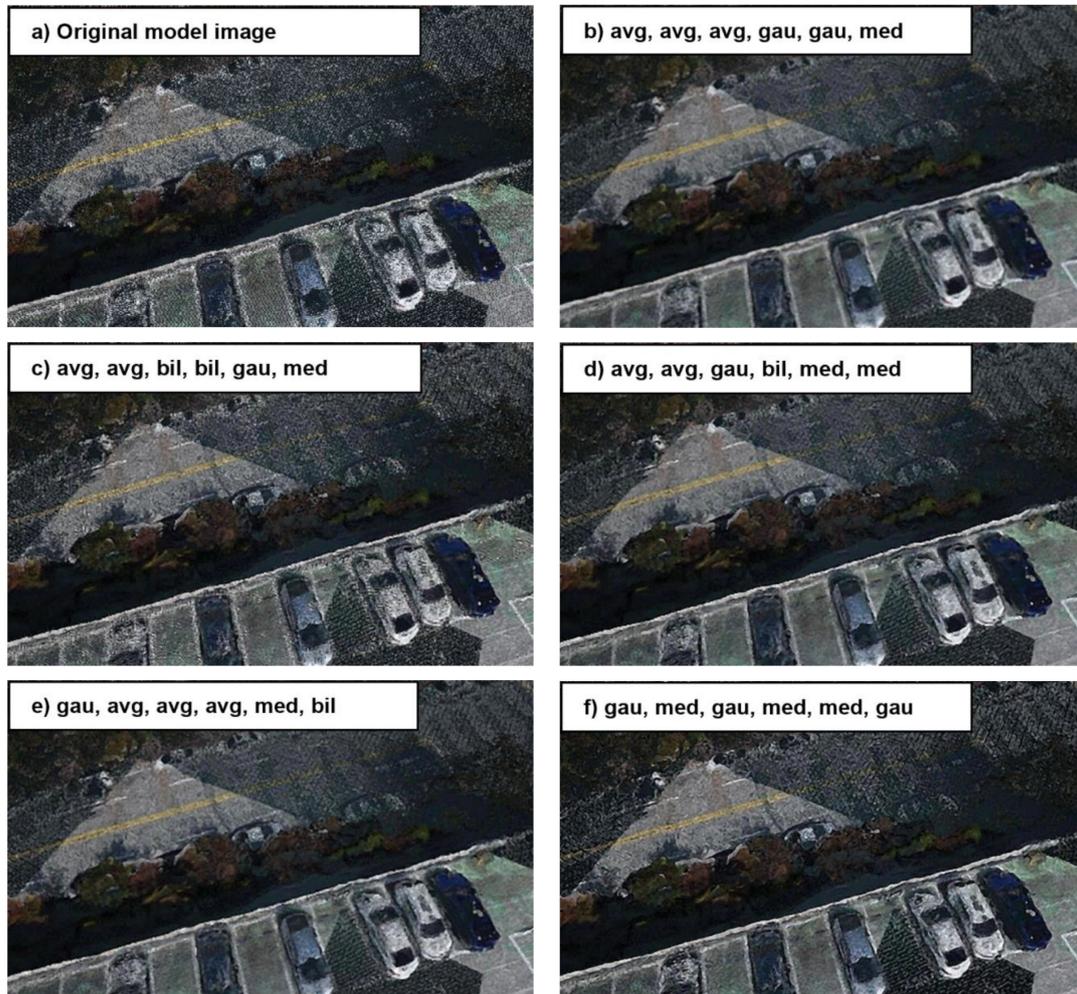
Figure 19. (**a**) The original image of PCM-based PCM, (**b**–**f**) exemplary filter combination. (**b**) MSE = 1146.46, PSNR = 17.54 dB. (**c**) MSE = 1103.24, PSNR = 17.70 dB. (**d**) MSE = 1199.40, PSNR = 17.34 dB, (**e**) MSE = 1472.52, PSNR = 16.45 dB. (**f**) MSE = 1622.87, PSNR = 16.03 dB.
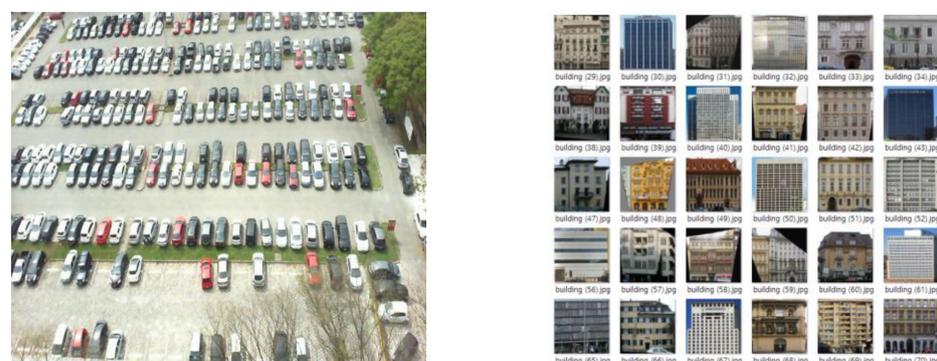
Figure 20. Car dataset (**left**); building dataset (**right**).

Site A:

Site B:

Site C:

Site D:
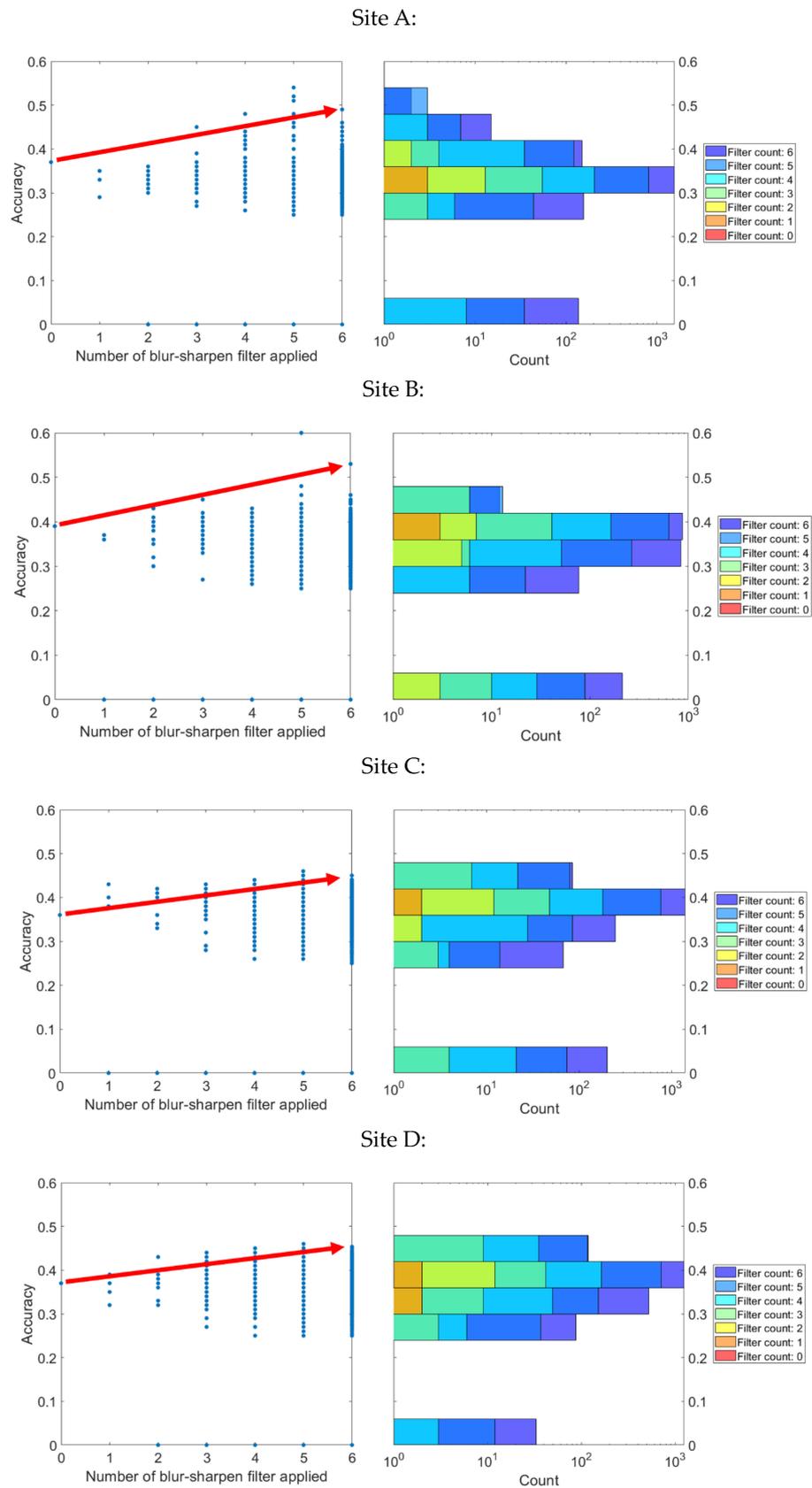
Figure 21. YOLO 20: YOLO object detection accuracy results based on the frequency of blur/sharpen filter applications. Four sites were evaluated, revealing specific combinations of blur/sharpen filter applications that enhanced the accuracy of object detection in images.

4.2. Generative Adversarial Networks (GAN)

Thus, to increase the PCM's photorealistic rendering resolution, we used a generative adversarial network (GAN) AI model to transform sketches into images. GAN models have the advantage of turning sketches into high-quality images, making it possible to obtain more refined photorealistic rendering of PCM models than the blur/sharpen filter if PCM models are defined as sketches and photorealistic rendering is performed.

The internal structure of GAN, as shown in Figure 22, is composed of a generator model and a discriminator model. The generator model aims to create random data, while the discriminator model aims to distinguish between actual and generated data. These two models compete, with the generator model trying to deceive the discriminator model and the discriminator model trying to differentiate the generated data from the actual data. In this way, the generator model can create data similar to actual data.
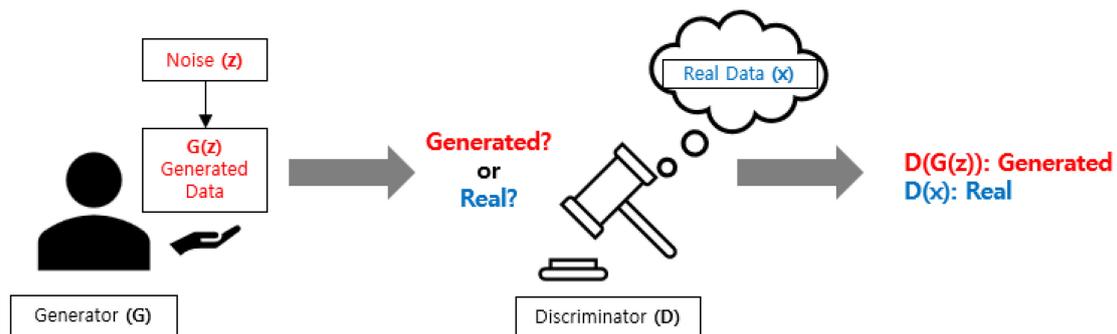

Figure 22. Structure of GAN AI model.

To use GAN, the training dataset must be composed of noisy and corresponding noise-free data. However, collecting or generating separate PCM training data for model photorealistic rendering contradicts one of the goals of this study, which is to minimize the training data. To solve this problem, we automatically generated a training dataset by adding point cloud noise to drone-captured images used to create the on-site PCM using WebODM, as shown in Figure 23.

Figure 23. GAN training dataset: drone-captured photo (**left**); photo after adding point cloud noise (**right**).

Furthermore, the versatility of the GAN AI model is enhanced by training it with construction site images captured under a variety of environmental conditions. This strategy enables the model to learn how to generate images where adverse environmental effects, which could otherwise impede object detection by acting as visual noise, are effectively minimized or eliminated. This approach ensures that the GAN AI model can produce images with reduced environmental noise, thereby facilitating more accurate object detection in varied site conditions.

After the training was completed, we rendered the PCM model using GAN, as shown in Figure 24's "After photorealistic rendering (middle)". Subsequently, the same YOLOv5 model used for object recognition after applying the blur/sharpen filter in Chapter 3 was used for object recognition in the photorealistic rendered PCM model.

Figure 24. Before photorealistic rendering (**left**); after photorealistic rendering (**middle**); object detection from photorealistic rendered PCM model (**right**).

As a result of object recognition, the detected objects increased in number and accuracy compared with when the blur/sharpen filter was applied. However, we observed that the resolution of the photorealistic rendered model varied significantly depending on the photo. The variation in resolution affected object recognition rates, necessitating additional recognition rates and accuracy for actual site management. GAN is one of the highest-performing AI models for converting sketches to images among existing models [40,41]. While achieving photorealistic rendering using a super-large AI model is possible, equipping such infrastructure on a construction site is practically challenging, and creating a more refined photorealistic model is complex.

4.3. Fine-Tuning of YOLO v5

After confirming the feasibility of the basic GAN model, we conducted a fine-tuning process on the YOLOv5 model for additional accuracy improvement. Fine-tuning refers to training an already trained model with a new dataset tailored to a specific situation. By adding a new training dataset based on the already trained model, the weights within the neural network are updated, improving performance. This method may contradict the research goal of reducing the effort required to acquire training data. However, fine-tuning can have a meaningful impact on accuracy even when using a small amount of high-quality data, such as 50 images [70]. In this study, we applied this fine-tuning technique to the YOLOv5 AI model.

For fine-tuning, additional training datasets were used, consisting of photos containing actual buildings and cars on-site for the purpose of recognizing buildings and cars. As a result of the tuning, the object recognition accuracy increased to (a) 48%, (b) 77%, (c) 100%,

and (d) 82%, as shown in Figure 25. Overall, it showed an increase of 48–82%, and it can be said that the recognition rate was successfully raised using the fine-tuning technique. However, there is a problem with no data on the equipment and buildings deployed on the initial construction site. This problem can be solved by synthesizing photos of objects to be introduced or placed on-site with actual site photos, as shown in studies [71,72], to create additional training datasets for fine-tuning.
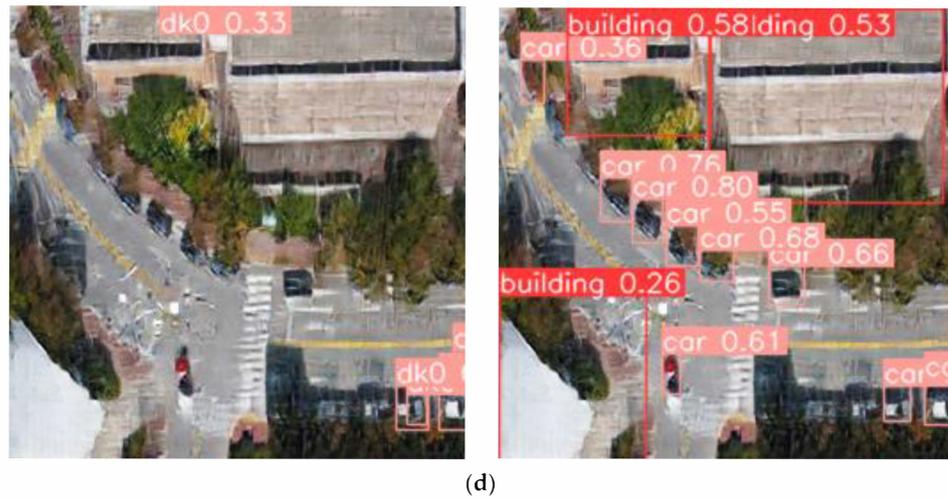

(a)

(b)

(c)

Figure 25. *Cont.*

(**d**)

Figure 25. Before fine-tuning (**left**); after fine-tuning (**right**).

4.4. Comparative Analysis of Results

The results from different steps of the research are summarized in Table 4 below for a better comparison. Initially, the blur/sharpen filter was used to make the model look more realistic, which helped the YOLOv5 AI, using a standard training dataset, to recognize objects in the model. However, to improve the image quality further, a generative adversarial network (GAN) sketch-to-image AI model was used instead of the blur/sharpen filter. This change resulted in better-quality images, which in turn improved object recognition accuracy.

Table 4. Results of object detection from each phase.

Phase	Explanation	Results
Blur/sharpen filter	- Photorealistic rendering of PCM-based PCM models was carried out by applying the blur/sharpen filter - The photorealistic rendered model reduced noise, allowing objects to be well recognized.	
GAN	- GAN, a sketch-to-image AI model, was used to achieve higher-resolution photorealistic rendering of PCM models. - The photorealistic rendered model showed a higher level of completion than the blur/sharpen filter.	
Fine-tuning	- The training dataset for the object recognition AI YOLOv5 model was changed to recognize objects in the PCM model rendered by GAN. - It showed the highest object recognition accuracy.	

Despite this improvement, the accuracy was still not good enough for on-site management. Therefore, a fine-tuning process was carried out to obtain better accuracy. For fine-tuning, real objects from the site were labeled to create a new training dataset. Through this fine-tuning process, the object recognition rate and accuracy were improved, even when the image quality was low. This step showed that fine-tuning could help in improving object recognition, which is crucial for better management of construction sites.

4.5. Limitations

This study encountered limitations beginning with the photorealistic rendering process. The application of blur/sharpen filters and GAN models for this purpose did not meet the expected success. These methods did not achieve the original goal of creating photorealistic rendering for object detection, indicating potential shortcomings in the choice of methods or their application.

Building on this issue, the constraints related to data and training further affected the study. The automatic generation of training data and minimal usage of additional training data may have limited the accuracy of the models. Specific construction site data, experimental design, and additional testing scenarios might have enhanced the results.

These technological challenges dovetail with the sensitivity of resolution and distance in the object recognition process. The effectiveness of recognition at close range sharply contrasted with decreased accuracy as distance increased. This divergence points to a need for further research to optimize recognition at varying distances.

Furthermore, the practical and scalability challenges of implementing high-quality rendering using advanced AI models pose barriers to application on construction sites. The methods may not be directly applicable or scalable to diverse environments, potentially limiting their broader use.

Finally, these limitations underscore the future research needs highlighted by the study. The necessity for developing more advanced rendering algorithms or integrating other technologies like LIDAR presents a pathway to address the current limitations and lead to more robust digital construction site management applications.

5. Conclusions and Future Research

This study highlights the potential of drone imagery in enhancing digital construction site management and Building Information Modeling (BIM) creation, particularly through improved object recognition within point cloud models (PCMs). A crucial finding is the need for precise adjustments in drone flight parameters—altitude, camera angles, and distances during image capture—to create accurate PCM models. The application of image preprocessing techniques, generative adversarial network (GAN) models for texture generation, and the refinement of object recognition algorithms with extensive real-world datasets has shown promise in enhancing the accuracy of object recognition. These advancements have significant implications for construction project management, including:

(1) Enhanced precision in project planning and monitoring: Accurate PCM models allow for detailed site analysis and monitoring, supporting informed decision-making based on precise, real-time data.

(2) Improved safety and risk management: Advanced object recognition capabilities can identify potential safety hazards and ensure compliance with safety protocols, thereby mitigating risks and enhancing onsite safety.

(3) Optimized resource allocation: Detailed insights into site conditions and progress from accurate digital models facilitate better resource allocation, reducing waste and increasing efficiency.

(4) Streamlined collaboration and communication: Digital models that accurately reflect the construction site condition improve communication among stakeholders, facilitating effective collaboration and coordination.

However, several limitations were encountered:

(a) The aim of achieving photorealistic rendering for object detection using blur/sharpen filters and GAN models was not fully met, indicating a need for alternative or refined methods.

(b) The effectiveness of object recognition varied with distance, suggesting further research is needed to optimize recognition at varying distances.

(c) Practical and scalability challenges emerged when attempting to implement high-quality rendering using advanced AI models on construction sites, indicating the methods may not be directly applicable or scalable to diverse environments.

These include challenges in achieving photorealistic rendering for object detection and the variable effectiveness of object recognition algorithms at different distances. Future research could explore the development of advanced rendering algorithms or the integration of technologies like LIDAR to overcome these challenges. Additionally, the automation of drone flight paths and image capture methods through AI-driven algorithms could further enhance the efficiency and accuracy of PCM model creation, significantly contributing to the digitalization of construction sites and improving productivity and innovation within the construction industry.

Author Contributions: S.H. was responsible for research planning, experimentation, and system identification; S.N. contributed to research writing, reference research, and analysis research; W.C., conducted experiments and assisted with organizing and analyzing data. Each author played a crucial role in the success of this study, and their contributions are greatly appreciated. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The datasets generated and/or analyzed during the current study are not publicly available due to confidentiality agreements with the construction companies involved in the study. This is to protect proprietary information and trade secrets pertaining to construction techniques, methodologies, and materials used, which could give these companies a competitive disadvantage if made public. However, they are available from the corresponding author upon reasonable request.

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

References

1.   Enshassi, M.A.; Al Hallaq, K.A.; Tayeh, B.A. Limitation factors of building information modeling (BIM) implementation. *Open Constr. Build. Technol. J.* **2019**, *13*, 189–196. [CrossRef]

2.   Sun, C.; Jiang, S.; Skibniewski, M.J.; Man, Q.; Shen, L. A literature review of the factors limiting the application of BIM in the construction industry. *Technol. Econ. Dev. Econ.* **2015**, *23*, 764–779. [CrossRef]

3.   Tang, S.; Shelden, D.R.; Eastman, C.M.; Pishdad-Bozorgi, P.; Gao, X. A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. *Autom. Constr.* **2019**, *101*, 127–139. [CrossRef]

4.   Azhar, S.; Khalfan, M.; Maqsood, T. Building information modelling (BIM): Now and beyond. *Constr. Econ. Build.* **2015**, *12*, 15–28. [CrossRef]

5.   Edirisinghe, R.J.E. Construction and A. Management, Digital skin of the construction site: Smart sensor technologies towards the future smart construction site. *Emerald Insight* **2019**, *26*, 184–223.

6.   Cabeza, L.F.; Rincón, L.; Vilariño, V.; Pérez, G.; Castell, A. Life cycle assessment (LCA) and life cycle energy analysis (LCEA) of buildings and the building sector: A review. *Renew. Sustain. Energy Rev.* **2014**, *29*, 394–416. [CrossRef]

7.   Wang, Q.; Kim, M.-K. Applications of 3D point cloud data in the construction industry: A fifteen-year review from 2004 to 2018. *Adv. Eng. Inform.* **2019**, *39*, 306–319. [CrossRef]

8.   Kavaliauskas, P.; Fernandez, J.B.; McGuinness, K.; Jurelionis, A. Automation of Construction Progress Monitoring by Integrating 3D Point Cloud Data with an IFC-Based BIM Model. *Buildings* **2022**, *12*, 1754. [CrossRef]

9.   Rebolj, D.; Pučko, Z.; Babič, N.Č.; Bizjak, M.; Mongus, D. Point cloud quality requirements for Scan-vs-BIM based automated construction progress monitoring. *Autom. Constr.* **2017**, *84*, 323–334. [CrossRef]

10. Kim, P.; Chen, J.; Cho, Y.K. SLAM-driven robotic mapping and registration of 3D point clouds. *Autom. Constr.* **2018**, *89*, 38–48. [CrossRef]

11. Kim, S.; Kim, S.; Lee, D.-E. 3D point cloud and BIM-based reconstruction for evaluation of project by as-planned and as-built. *Remote Sens.* **2020**, *12*, 1457. [CrossRef]

12. Melenbrink, N.; Werfel, J.; Menges, A. On-site autonomous construction robots: Towards unsupervised building. *Autom. Constr.* **2020**, *119*, 103312. [CrossRef]

13. Delgado, J.M.D.; Oyedele, L.; Ajayi, A.; Akanbi, L.; Akinade, O.; Bilal, M.; Owolabi, H. Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *J. Build. Eng.* **2019**, *26*, 100868. [CrossRef]

14. Huang, Z.; Mao, C.; Wang, J.; Sadick, A.-M. Understanding the key takeaway of construction robots towards construction automation. *Eng. Constr. Arch. Manag.* **2021**, *29*, 3664–3688. [CrossRef]

15. Heo, S.; Na, S.; Han, S.; Shin, Y.; Lee, S. Flip side of artificial intelligence technologies: New labor-intensive industry of the 21st century. *J. Comput. Struct. Eng. Inst. Korea* **2021**, *34*, 327–337. [CrossRef]

16. Jiang, Y. Intelligent building construction management based on BIM digital twin. *Comput. Intell. Neurosci.* **2021**, *2021*, 4979249. [CrossRef] [PubMed]

17. Parusheva, S. Business, and Education, Digitalization and Digital Transformation in Construction-Benefits and Challenges. In *Information and Communication Technologies in Business and Education*; University of Economics: Varna, Bulgaria, 2019; pp. 126–134.

18. Lundberg, O.; Nylén, D.; Sandberg, J. Unpacking construction site digitalization: The role of incongruence and inconsistency in technological frames. *Constr. Manag. Econ.* **2022**, *40*, 987–1002. [CrossRef]

19. Nti, I.K.; Adekoya, A.F.; Weyori, B.A.; Nyarko-Boateng, O. Applications of artificial intelligence in engineering and manufacturing: A systematic review. *J. Intell. Manuf.* **2022**, *33*, 1581–1601. [CrossRef]

20. Asadi, K.; Suresh, A.K.; Ender, A.; Gotad, S.; Maniyar, S.; Anand, S.; Noghabaei, M.; Han, K.; Lobaton, E.; Wu, T. An integrated UGV-UAV system for construction site data collection. *Autom. Constr.* **2020**, *112*, 103068. [CrossRef]

21. Rachmawati, T.S.N.; Kim, S. Unmanned aerial vehicles (UAV) integration with digital technologies toward construction 4.0: A systematic literature review. *Sustainability* **2022**, *14*, 5708. [CrossRef]

22. Coupry, C.; Noblecourt, S.; Richard, P.; Baudry, D.; Bigaud, D. BIM-based digital twin and xr devices to improve maintenance procedures in smart buildings: A literature review. *Appl. Sci.* **2022**, *11*, 6810. [CrossRef]

23. Opoku, D.-G.J.; Perera, S.; Osei-Kyei, R.; Rashidi, M. Digital twin application in the construction industry: A literature review. *J. Build. Eng.* **2021**, *40*, 102726. [CrossRef]

24. Semeraro, C.; Lezoche, M.; Panetto, H.; Dassisti, M. Digital twin paradigm: A systematic literature review. *Comput. Ind.* **2021**, *130*, 103469. [CrossRef]

25. Pan, Y.; Zhang, L. A BIM-data mining integrated digital twin framework for advanced project management. *Autom. Constr.* **2021**, *124*, 103564. [CrossRef]

26. Sacks, R.; Brilakis, I.; Pikas, E.; Xie, H.S.; Girolami, M. Construction with digital twin information systems. *Data-Centric Eng.* **2020**, *1*, e14. [CrossRef]

27. Xue, H.; Zhang, S.; Su, Y.; Wu, Z.; Yang, R.J. Effect of stakeholder collaborative management on off-site construction cost performance. *J. Clean. Prod.* **2018**, *184*, 490–502. [CrossRef]

28. Razkenari, M.; Fenner, A.; Shojaei, A.; Hakim, H.; Kibert, C. Perceptions of offsite construction in the United States: An investigation of current practices. *J. Build. Eng.* **2019**, *29*, 101138. [CrossRef]

29. Abioye, S.O.; Oyedele, L.O.; Akanbi, L.; Ajayi, A.; Delgado, J.M.D.; Bilal, M.; Akinade, O.O.; Ahmed, A. Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *J. Build. Eng.* **2021**, *44*, 103299. [CrossRef]

30. Wu, J.; Zheng, H.; Zhao, B.; Li, Y.; Yan, B.; Liang, R.; Wang, W.; Zhou, S.; Lin, G.; Fu, Y.J. Ai challenger: A large-scale dataset for going deeper in image understanding. *arXiv* **2017**, arXiv:1711.06475.

31. Zhou, W.; Whyte, J.; Sacks, R. Construction safety and digital design: A review. *Autom. Constr.* **2012**, *22*, 102–111. [CrossRef]

32. Afzal, M.; Shafiq, M.T.; Al Jassmi, H. Improving construction safety with virtual-design construction technologies—A review. *J. Inf. Technol. Constr.* **2021**, *26*, 319–340. [CrossRef]

33. Torrecilla-García, J.A.; Pardo-Ferreira, M.C.; Rubio-Romero, J.C. Overall Introduction to the Framework of BIM-based Digital Twinning in Decision-making in Safety Management in Building Construction Industry. *Dir. Organ.* **2021**, *74*, 31–38. [CrossRef]

34. Collinge, W.H.; Farghaly, K.; Mosleh, M.H.; Manu, P.; Cheung, C.M.; Osorio-Sandoval, C.A. BIM-based construction safety risk library. *Autom. Constr.* **2022**, *141*, 104391. [CrossRef]

35. Guo, H.; Scheepbouwer, E.; Yiu, T.; Gonzalez, V. *Overview and Analysis of Digital Technologies for Construction Safety Management*; University of Canberra: Canberra, Australia, 2017.

36. Parsamehr, M.; Perera, U.S.; Dodanwala, T.C.; Perera, P.; Ruparathna, R. A review of construction management challenges and BIM-based solutions: Perspectives from the schedule, cost, quality, and safety management. *Asian J. Civ. Eng.* **2023**, *24*, 353–389. [CrossRef]

37. Nakanishi, Y.; Kaneta, T.; Nishino, S. A Review of Monitoring Construction Equipment in Support of Construction Project Management. *Front. Built Environ.* **2022**, *7*, e0632593. [CrossRef]

38. Forteza, F.J.; Carretero-Gómez, J.M.; Sesé, A. Safety in the construction industry: Accidents and precursors. *J. Constr.* **2020**, *19*, 271–281. [CrossRef]

39. Kang, Y.; Siddiqui, S.; Suk, S.J.; Chi, S.; Kim, C. Trends of Fall Accidents in the U.S. Construction Industry. *J. Constr. Eng. Manag.* **2017**, *143*, e0001332. [CrossRef]

40. Follini, C.; Magnago, V.; Freitag, K.; Terzer, M.; Marcher, C.; Riedl, M.; Giusti, A.; Matt, D.T. BIM-integrated collaborative robotics for application in building construction and maintenance. *Robotics* **2020**, *10*, 2. [CrossRef]

41. Torres-González, M.; Prieto, A.; Alejandre, F.; Blasco-López, F. Digital management focused on the preventive maintenance of World Heritage Sites. *Autom. Constr.* **2021**, *129*, 103813. [CrossRef]

42. Errandonea, I.; Beltrán, S.; Arrizabalaga, S. Digital Twin for maintenance: A literature review. *Comput. Ind.* **2020**, *123*, 103316. [CrossRef]

43. Rødseth, H.; Schjølberg, P.; Marhaug, A. Deep digital maintenance. *Adv. Manuf.* **2017**, *5*, 299–310. [CrossRef]

44. Lu, Q.; Xie, X.; Parlikad, A.K.; Schooling, J.M. Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. *Autom. Constr.* **2020**, *118*, 103277. [CrossRef]

45. Zavadskas, E.K. Automation and robotics in construction: International research and achievements. *Autom. Constr.* **2010**, *19*, 286–290. [CrossRef]

46. Bogue, R. What are the prospects for robots in the construction industry? *Ind. Robot. Int. J. Robot. Res. Appl.* **2018**, *45*, 1–6. [CrossRef]

47. Na, S.; Heo, S.; Han, S.; Shin, Y.; Roh, Y. Acceptance model of artificial intelligence (AI)-based technologies in construction firms: Applying the technology acceptance model (tam) in combination with the technology–organisation–environment (TOE) framework. *Buildings* **2022**, *12*, 90. [CrossRef]

48. Chen, X.; Chang-Richards, A.Y.; Pelosi, A.; Jia, Y.; Shen, X.; Siddiqui, M.K.; Yang, N. Implementation of technologies in the construction industry: A systematic review. *Eng. Constr. Arch. Manag.* **2022**, *29*, 3181–3209. [CrossRef]

49. Olanipekun, A.O.; Sutrisna, M. Facilitating digital transformation in construction—A systematic review of the current state of the art. *Front. Built Environ.* **2021**, *7*, 660758. [CrossRef]

50. Zhou, W.; Georgakis, P.; Heesom, D.; Feng, X. Model-based groupware solution for distributed real-time collaborative 4D planning through teamwork. *J. Comput. Civ. Eng.* **2012**, *26*, 597–611. [CrossRef]

51. Zhang, S.; Teizer, J.; Lee, J.-K.; Eastman, C.M.; Venugopal, M. Building information modeling (BIM) and safety: Automatic safety checking of construction models and schedules. *Autom. Constr.* **2013**, *29*, 183–195. [CrossRef]

52. Levin, D.I.W.; Litven, J.; Jones, G.L.; Sueda, S.; Pai, D.K. Eulerian solid simulation with contact. *ACM Trans. Graph.* **2011**, *30*, 1–10. [CrossRef]

53. García, J.M.B. Recording stratigraphic relationships among non-original deposits on a 16th century painting. *J. Cult. Herit.* **2009**, *10*, 338–346. [CrossRef]

54. Cuypers, S.; Bassier, M.; Vergauwen, M. Deep Learning on Construction Sites: A Case Study of Sparse Data Learning Techniques for Rebar Segmentation. *Sensors* **2021**, *21*, 5428. [CrossRef] [PubMed]

55. Akinosho, T.D.; Oyedele, L.O.; Bilal, M.; Ajayi, A.O.; Delgado, M.D.; Akinade, O.O.; Ahmed, A.A. Deep learning in the construction industry: A review of present status and future innovations. *J. Build. Eng.* **2020**, *32*, 101827. [CrossRef]

56. Lee, J.; Lee, S. Construction site safety management: A computer vision and deep learning approach. *Sensors* **2023**, *23*, 944. [CrossRef] [PubMed]

57. Nath, N.D.; Behzadan, A.H. Deep convolutional networks for construction object detection under different visual conditions. *Front. Built Environ.* **2020**, *6*, 97. [CrossRef]

58. Tang, S.; Roberts, D.; Golparvar-Fard, M. Human-object interaction recognition for automatic construction site safety inspection. *Autom. Constr.* **2020**, *120*, 103356. [CrossRef]

59. Muhammad, I.; Ying, K.; Nithish, M.; Xin, J.; Xinge, Z.; Cheah, C.C. Robot-assisted object detection for construction automation: Data and information-driven approach. *IEEE/ASME Trans. Mechatron.* **2021**, *26*, 2845–2856. [CrossRef]

60. Heo, S.; Han, S.; Shin, Y.; Na, S. Challenges of data refining process during the artificial intelligence development projects in the architecture, engineering and construction industry. *Appl. Sci.* **2021**, *11*, 10919. [CrossRef]

61. Shin, Y.; Heo, S.; Han, S.; Kim, J.; Na, S. An image-based steel rebar size estimation and counting method using a convolutional neural network combined with homography. *Buildings* **2021**, *11*, 463. [CrossRef]

62. Sunwoo, H.; Choi, W.; Na, S.; Kim, C.; Heo, S. Comparison of the Performance of Artificial Intelligence Models Depending on the Labelled Image by Different User Levels. *Appl. Sci.* **2022**, *12*, 3136. [CrossRef]

63. Lee, M. Recent Advances in Generative Adversarial Networks for Gene Expression Data: A Comprehensive Review. *Mathematics* **2023**, *11*, 3055. [CrossRef]

64. Terven, J.; Cordova-Esparza, D. A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. *arXiv* **2023**, arXiv:2304.00501.

65. Gašparović, B.; Mauša, G.; Rukavina, J.; Lerga, J. Evaluating YOLOV5, YOLOV6, YOLOV7, and YOLOV8 in Underwater Environment: Is There Real Improvement? In Proceedings of the 8th International Conference on Smart and Sustainable Technologies (SpliTech), Split/Bol, Croatia, 1 August 2023; pp. 1–4.

66. Halder, S.; Afsari, K.; Serdakowski, J.; DeVito, S.; Ensafi, M.; Thabet, W. Real-Time and Remote Construction Progress Monitoring with a Quadruped Robot Using Augmented Reality. *Buildings* **2022**, *12*, 2027. [CrossRef]

67. Bassier, M.; Vermandere, J.; De Winter, H. Linked building data for construction site monitoring: A test case. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *V-2-2022*, 159–165. [CrossRef]

68. Universidade Federal do Parana. PKLot Dataset. In *Roboflow Universe*; Universidade Federal do Parana: Curitiba, Brazil, 2022; Available online: https://public.roboflow.com/object-detection/pklot (accessed on 10 April 2024).

69. Istanbul Technical University. facadesArchitecture Dataset. In *Roboflow Universe*; Istanbul Technical University: Istanbul, Turkey, 2022; Available online: https://universe.roboflow.com/stanbul-technical-university/facadesarchitecture (accessed on 10 April 2024).

70. Zhao, M.; Abbeel, P.; James, S. On the effectiveness of fine-tuning versus meta-reinforcement learning. *NeurIPS* **2022**, *35*, 26519–26531.

71. Bhowmik, N.; Wang, Q.; Gaus, Y.F.A.; Szarek, M.; Breckon, T.P.J. The good, the bad and the ugly: Evaluating convolutional neural networks for prohibited item detection using real and synthetically composed X-ray imagery. *arXiv* **2019**, arXiv:1909.11508.

72. Jeon, M.; Student, R.P.M.; Lee, Y.; Shin, Y.-S.; Jang, H.; Yeu, T.; Kim, A. Synthesizing Image and Automated Annotation Tool for CNN based Under Water Object Detection. *J. Korea Robot. Soc.* **2019**, *14*, 139–149. [CrossRef]