buildings

MDPI

Article

Innovative Point Cloud Segmentation of 3D Light Steel Framing System through Synthetic BIM and Mixed Reality Data: Advancing Construction Monitoring

Yee Sye Lee [1], Ali Rashidi [1,2,*], Amin Talei [1,3] and Daniel Kong [1,3]

[1] Department of Civil Engineering, School of Engineering, Monash University Malaysia, Bandar Sunway 47500, Malaysia
[2] Department of Engineering, La Trobe University, Melbourne, VIC 3086, Australia
[3] Monash Climate-Resilient Infrastructure Research Hub (M-CRInfra), School of Engineering, Monash University Malaysia, Bandar Sunway 47500, Malaysia
* Correspondence: a.rashidi@latrobe.edu.au

Abstract: In recent years, mixed reality (MR) technology has gained popularity in construction management due to its real-time visualisation capability to facilitate on-site decision-making tasks. The semantic segmentation of building components provides an attractive solution towards digital construction monitoring, reducing workloads through automation techniques. Nevertheless, data shortages remain an issue in maximizing the performance potential of deep learning segmentation methods. The primary aim of this study is to address this issue through synthetic data generation using Building Information Modelling (BIM) models. This study presents a point-cloud-based deep learning segmentation approach to a 3D light steel framing (LSF) system through synthetic BIM models and as-built data captured using MR headsets. A standardisation workflow between BIM and MR models was introduced to enable seamless data exchange across both domains. A total of five different experiments were set up to identify the benefits of synthetic BIM data in supplementing actual as-built data for model training. The results showed that the average testing accuracy using solely as-built data stood at 82.88%. Meanwhile, the introduction of synthetic BIM data into the training dataset led to an improved testing accuracy of 86.15%. A hybrid dataset also enabled the model to segment both the BIM and as-built data captured using an MR headset at an average accuracy of 79.55%. These findings indicate that synthetic BIM data have the potential to supplement actual data, reducing the costs associated with data acquisition. In addition, this study demonstrates that deep learning has the potential to automate construction monitoring tasks, aiding in the digitization of the construction industry.

Keywords: deep learning; mixed reality; construction monitoring; synthetic data; building information modelling

check for updates

Citation: Lee, Y.S.; Rashidi, A.; Talei, A.; Kong, D. Innovative Point Cloud Segmentation of 3D Light Steel Framing System through Synthetic BIM and Mixed Reality Data: Advancing Construction Monitoring. *Buildings* **2024**, *14*, 952. https://doi.org/10.3390/buildings14040952

Academic Editor: Paulo Santos

Received: 4 February 2024
Revised: 14 March 2024
Accepted: 26 March 2024
Published: 30 March 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Progress monitoring and controlling are the major project management tasks carried out throughout a construction project. These tasks are the foundation for ensuring that construction projects are completed on time and that economic interests are guaranteed [1]. The on-site activities involved in progress monitoring and controlling include the supervision of construction work, quantifying the level of completion of the work, assessing building quality, reporting on construction progress, and addressing on-site issues. Most progress monitoring and controlling tasks are still traditionally conducted using 2D drawings, reports, schedules, photo logs, and paper-based assembly instructions [2]. These pose problems, which are attributed to poor-quality management control due to their complexity and inefficient nature. However, the ability to capture and document as-built site information enables the generation of digital twins of ongoing construction, facilitating timely

inspection and control processes [3]. In this regard, the MR headset poses a compelling solution to enhance construction monitoring productivity while effectively reducing the risks associated with inadequate inspection practices.

The visual scene understanding of construction sites is a critical factor in the automated recognition of as-built data, aiding in the generation of digital twins. However, current practices tend to position cameras in a static manner, typically from one or multiple vantage points [4]. These stationary setups capture limited perspectives and angles, providing a partial view of the entire site. This approach often results in blind spots or areas that are not fully covered, potentially missing crucial details or progress updates in occluded areas. Meanwhile, MR headsets offer a dynamic solution by being mobile and adaptable to various viewpoints across a construction site. Unlike fixed cameras, these headsets allow users to move freely, exploring different angles and locations. This mobility enables on-site personnel to capture and monitor construction progress from multiple perspectives, providing a more holistic and real-time understanding of the construction process compared to static camera setups. While deep learning networks have proven to be helpful in the automated tasks in construction projects, the lack of sufficient training data poses significant challenges to their training process [5]. The acquisition of large and diverse datasets can be particularly challenging in emerging industries, such as using LSF systems in industrialised construction. This limitation often leads to overfitting or underfitting, where the model fails to generalize patterns and reliably perform its tasks. Addressing these issues requires methods such as data augmentation [6] or synthetic data generation [7] to supplement limited datasets, improving the robustness and effectiveness of deep learning models. In addition, most classification and regression approaches utilise input sources from which it is difficult to extract the geometrical or dimensional information of the whole structure or building [8,9]. Therefore, a point-cloud-based approach could potentially be further developed with spatial awareness for automated inspection in the future.

The authors present an approach that addresses this issue by generating synthetic data from BIM models to supplement the neural network training process. As the cost of data acquisition using BIM models is significantly lower than on-site data collection, this approach provides an attractive alternative that can be conducted on a larger scale. Furthermore, using MR headsets allows for the real-time validation of the as-built environment, highlighting any discrepancies in the BIM models against it and enabling prompt corrective action. This research aims to leverage BIM data to supplement the on-site data collection of LSF systems obtained through MR headsets. Therefore, our research objectives can be formulated as follows:

i. Assess the efficacy of synthetic data generated by BIM models in enhancing a neural network's performance in the point-cloud-based segmentation of LSF systems.
ii. Assess the hybrid effects of using synthetic data and as-built scans obtained directly from MR headsets in actual construction scenarios.

2. Mixed Reality and Deep Learning Approaches in the Construction Industry

Milgram and Kishino defined MR as a "reality spectrum", ranging from reality to virtual reality [10]. The first usage of MR technology in the construction industry can be found in a study conducted by Feiner [11], for overlaying information on a building. Since then, it has been widely used in visualizing 3D building models to facilitate construction designs [12], identify construction clashes [13], facilitate facilities management [14], and so on. This technology also enables effective communication, enhancing the safety communication on construction sites [15]. In addition, this technology can potentially guide users towards better decision making and learning skills while on construction sites [16,17]. This includes training crane operators [18] and other construction equipment operators [19], demonstrating its potential as a training platform. The surge of BIM's adoption has also led to an opportunity to improve modular and industrialised construction [20]. However, several risks are associated with adopting industrialised construction, including the inefficiency of the verification process of precast components [21]. This has led to a lack of

confidence in demonstrating industrialised construction's practicability and scalability, which could be resolved through automation techniques [22].

Deep learning has emerged as a transformative force in diverse areas of the construction industry, including construction safety, progress monitoring, inspection, and so on [23]. A vision-based system was developed to inspect steel frame assemblies off-site using edge detection methods [24]. A Stacked Auto-Encoders (SAEs) technique was also developed to classify 3D models in a BIM environment [25]. Some researchers have also utilised empirical wavelet transforms (EWTs) to identify structural damage [26]. Deep learning holds promise in industrialised construction by optimizing production processes through predictive analytics and automated quality control measures, streamlining efficiency and precision [27]. Nevertheless, the cost associated with training deep learning networks to provide an adequate level of accuracy remains one of the issues in model training [28]. There are a few proposed solutions, one of which involves data augmentation, essentially applying transformations such as translations and rotations to generate additional training data [29,30]. In addition, synthetic data could potentially be used to accelerate data access without actual data, offering a cost-effective and scalable means to augment limited datasets [31,32].

Several studies have identified the use of MR and deep learning in industrialised construction projects. However, there are some limitations associated with its implementation, which are presented below in Table 1.

Table 1. Existing studies illustrating research gaps.

Reference	Brief Description	Cons *	
		A	B
Chalhoub and Ayer, 2018 [33]	Using MR to superimpose BIM models for visualization purposes	X	-
Riexinger et al., 2018 [34]	Using MR for on-site manual inspection using BIM models	X	-
Ren et al., 2023 [35]	2D vision-based method for concrete inspection	X	X
Truong et al., 2023 [36]	Hardwood flooring defect inspection using multiple cameras	X	X

* Note A: lack of automated interpretation of 3D BIM models; B: lack of in-depth information extraction capabilities in models.

Currently, most MR applications in the construction industry rely on the manual interpretation of 3D BIM models for installation, inspection, and monitoring tasks. Meanwhile, most automated inspection applications utilise 2D image inputs, from which it is difficult to extract geometrical or dimensional properties. As such, the combined use of synthetic and real data for segmenting industrialised building components presents us with an opportunity to enhance the construction monitoring process by providing comprehensive and diverse datasets. Not only that, but a point-cloud-based approach using MR headsets could be further developed to visualise the geometrical and dimensional discrepancies between as-built and as-planned data directly on site. This could potentially foster greater confidence in the reliability of automated monitoring systems, contributing to a more widespread acceptance and adoption of industrialised construction.

3. Methodology

This section is separated into three subsections, highlighting the processes involved in different stages of the research. Figure 1 shows the overall research design process and is explained step-by-step in the subsections.

Both synthetic BIM and as-built MR-captured data are used in the neural network training process. Different training scenarios are studied to understand the usefulness of synthetic data in complementing a lack of actual training data.

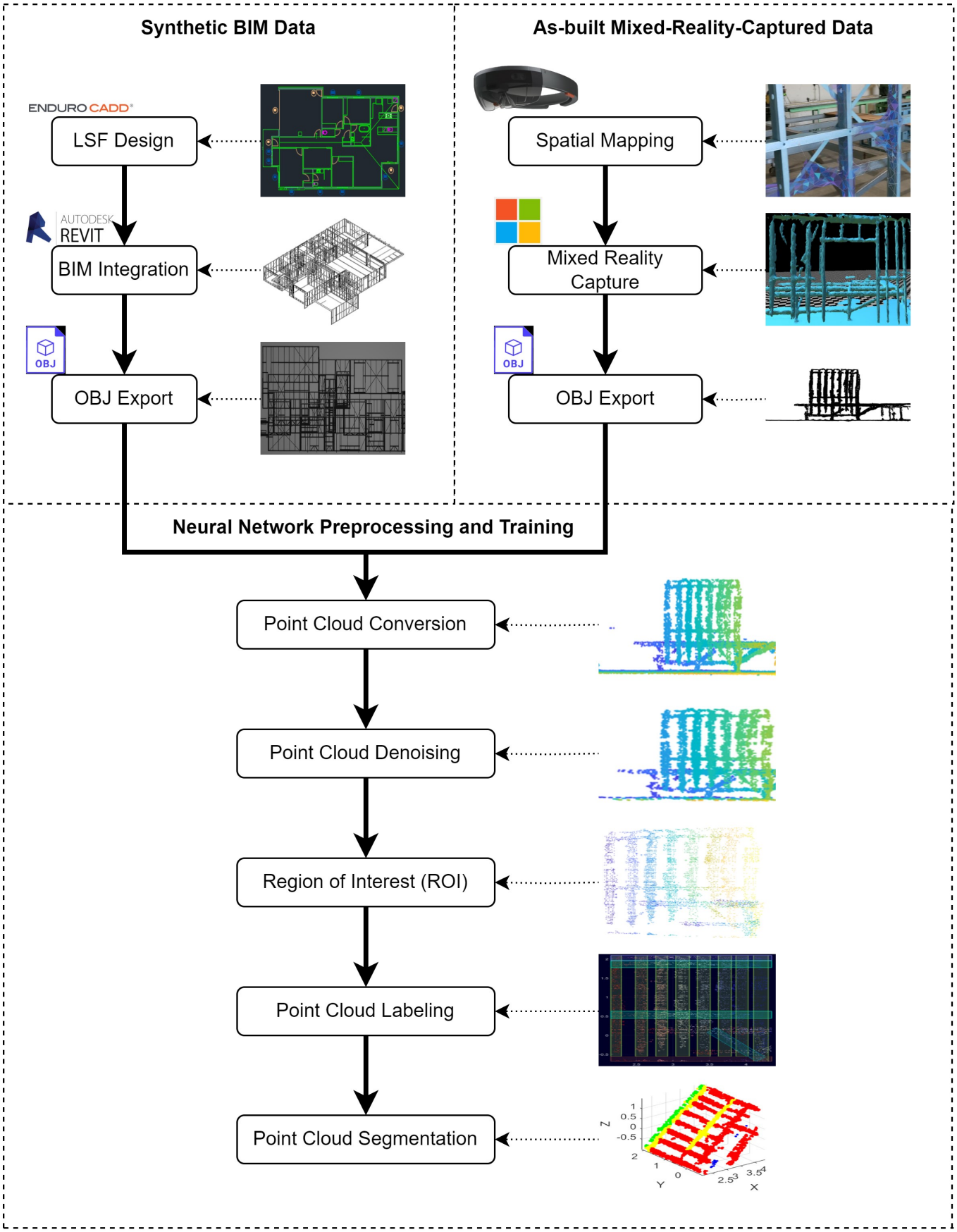


Figure 1. Schematic demonstration of the research design process.

3.1. Synthetic BIM Data

The process of obtaining synthetic BIM data for an LSF system is an initial phase in generating a dataset for neural network training. The synthetic LSF data are obtained directly from ENDUROCADD®v11, which is a steel frame software used to detail and engineer light-gauge steel house frames for manufacture using the ENDUROFRAME® rollformer, sourced by Steer Manufacturing in Victoria, Australia. The designs are exported as IFC files into Autodesk® Revit® (version 2023, San Francisco, CA, USA), a BIM software. Considering data availability, following the project's scope in real life, this research solely focuses on wall frames. This targeted approach simplifies the data acquisition process, reducing the number of frames needed to achieve representative data for training and testing at a later stage. To achieve this, a customized visibility filter was applied to obscure irrelevant structural objects, such as the flooring and roofing systems. A total of 10 wall frames were then selected from the BIM software, each consisting of diversified types of structural elements, including studs, top tracks, bottom tracks, noggins, and bracing. The C-channels used in the design are 90 mm deep and 0.75 mm or 1.00 mm thick. Meanwhile, the U-channels used are 93.5 mm in depth with a 0.75 mm thickness. The frames have a maximum height and span limit of 2600 mm and 5000 mm, respectively. A mixture of frames spanning the Y–Z and X–Z directions was used to account for directionality considerations at the training stage. An example of synthetic BIM data is shown in Figure 2a. After the preparation of the synthetic BIM dataset, it was then exported as an OBJ file for neural network training.

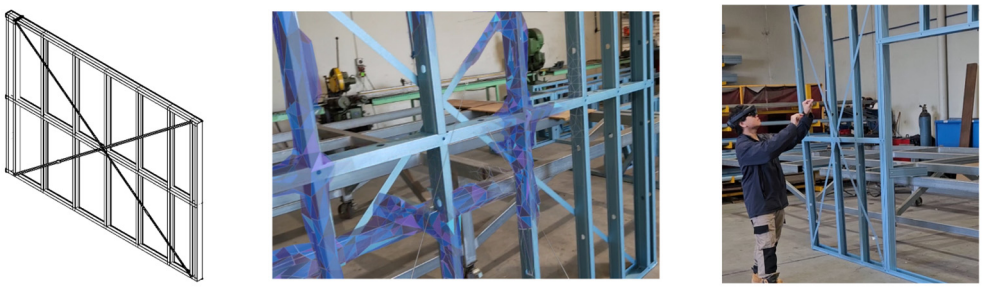


Figure 2. LSF training data's generation and capture: (**a**) example of LSF BIM data; (**b**) first-person view directly through a Microsoft HoloLens 2; (**c**) third-person view during the data collection process.

3.2. As-Built Mixed-Reality-Captured Data

The data acquisition of the as-built mixed-reality-captured data was conducted in a steel fabrication factory, Steer Manufacturing, which manufactures frames using the ENDUROFRAME® LSF system. The spatial mapping of the LSF system was conducted using the long-throw, low-frequency depth camera within the Microsoft HoloLens 2, an MR headset with a Qualcomm Snapdragon 850 SoC and 4 GB RAM. The HoloLens Research Mode sensor stream data were accessed through the Windows Device Portal on a laptop with the following specifications: AMD Ryzen 9 5900HX CPU, RTX 3070 GPU, and 16 GB RAM. The mesh caching of the spatial surfaces of the LSF system was captured on a 360-degree walkthrough of the object of interest, which encompassed detailed views from every angle to form a complete mesh containing each element of the LSF wall frames. Each spatial surface represents a partial volume of the LSF system, expressed as triangle meshes and multiple volumes from a complete view of the system. A total of 10 different wall frame data inputs containing a wide range of arrangements of their structural elements were obtained through this process. Similar to the synthetic data, the LSF wall frames contain a mixture of studs, top tracks, bottom tracks, noggins, and bracing elements. Nevertheless, the data arrangement of the as-built and synthetic data are set differently to prevent any generalization issues in recognizing unfamiliar patterns during neural network training. During data collection, each wall frame was propped vertically in an upright position and

fastened with bolts that connected it to a secured steel frame. The factory environment is equipped with adequate lighting, and the frames were located indoors. Figure 2b,c show an example of the first-person and third-person points of view during the capture process. After querying the mesh representation of the LSF system, it was then saved locally in an OBJ file format.

3.3. Neural Network Pre-Processing and Training

After the data collection process has been completed and exported as an OBJ file format, it is then imported into MATLAB® R2022b for further processing. Converting 3D objects into a standardized format for neural network training is crucial due to the myriad file formats used in different devices, such as PCD, PLY, LAS, LAZ, OBJ, FBX, etc. In this study, we opted to convert 3D objects into a standardized PCD format widely used and supported within robotics and autonomous systems. This ensures that the trained neural network is applicable to other sensors, such as LiDAR cameras, in the future. Firstly, the OBJ file is read and parsed to extract information about vertices, normal, and faces. Then, a point cloud is constructed using this information. Figure 3a,b show an example of an OBJ 3D object and the point cloud representation of that 3D object, respectively. As the data obtained from the MR headset often contain noises such as hallucinations, denoising techniques help reduce the irrelevant information by filtering out any outliers from the point cloud. In this study, for every point in the dataset, four neighbouring points are evaluated and considered to be outliers if the average distance to their nearest neighbours exceeds 0.02. An example of removing outliers can be found in Figure 3c, highlighting the removal of several points that are floating around. Afterwards, a 3D bounding box is formed to determine the region of interest (ROI), expressed as the minimum and maximum of the x,y, and z coordinates. The points within the specified ROI are obtained, indexed, and selected using a Kd-tree-based search algorithm, as shown in Figure 3d. Then, they are written into PCD file format using ASCII encoding.

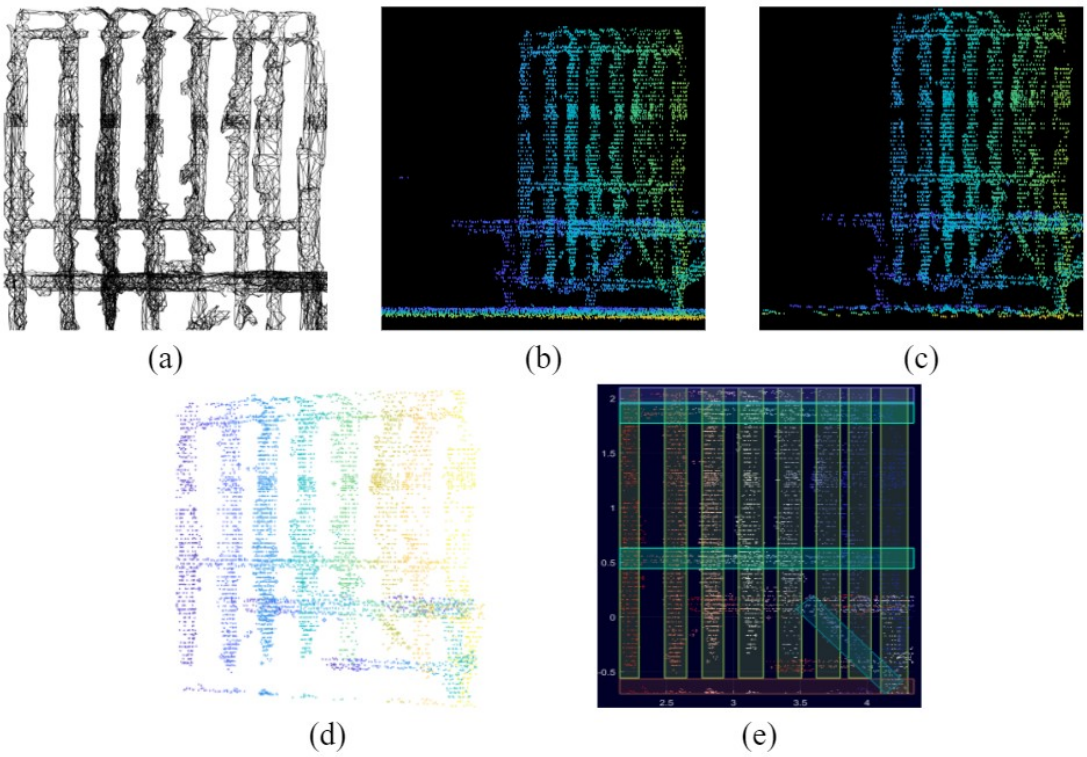


Figure 3. Neural network pre-processing and training: (**a**) example of a 3D Mesh OBJ; (**b**) example of a point cloud representation of an LSF system; (**c**) point cloud representation after removing outliers; (**d**) point cloud region of interest selection; (**e**) point cloud labelling.

The PCD file is used in point cloud labelling, annotating semantic labels to individual points using cuboid clusters. The points are assigned to either one of the following labels: unassigned, stud, top plate, bottom plate, noggin, or bracing. An example of the point cloud labelling process can be found in Figure 3e.

We then identify the points falling within the boundaries using the cuboids' parameters: $x_{ctr}, y_{ctr}, z_{ctr}$ represent the centre coordinates, $x_{len}, y_{len}, z_{len}$ represent the dimensions of the cuboids, and $\theta_x, \theta_y, \theta_z$ represent the rotation angles. The rotation matrices R_x, R_y, R_z can then be calculated using Equations (1)–(3), as follows:

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{bmatrix} \quad (1)$$

$$R_y = \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix} \quad (2)$$

$$R_z = \begin{bmatrix} \cos(\theta_z) & -\sin(\theta_z) & 0 \\ \sin(\theta_z) & \cos(\theta_z) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

The combined rotation matrix, R, is then computed by multiplying the rotation matrices R_x, R_y, R_z. The original point cloud locations, P_{ori}, are then translated and rotated relative to the cuboids' centre, P_{rel}, using Equation (4).

$$P_{rel} = R^{-1}[P_{ori} - [x_{ctr} \quad y_{ctr} \quad z_{ctr}]] \quad (4)$$

Then, a mask, m, is created across all axes using Equation (5).

$$m = \left(-\frac{x_{len}}{2} < P_{rel,x} < \frac{x_{len}}{2}\right) \Lambda \left(-\frac{y_{len}}{2} < P_{rel,y} < \frac{y_{len}}{2}\right) \Lambda \left(-\frac{z_{len}}{2} < P_{rel,z} < \frac{z_{len}}{2}\right) \quad (5)$$

The points that satisfy these conditions in all three dimensions indicate that they fall within the specified cuboid's volume in 3D space, and labels are assigned accordingly.

After the points have been assigned labels, the points are separated into three subsets: the training, validation, and testing datasets. The combination of synthetic data and as-built mixed-reality-captured data totals to 92,319 points after pre-processing. To investigate the effects of synthetic, as-built, and hybrid datasets, we established a total of five distinct experiments, as shown in Figure 4.

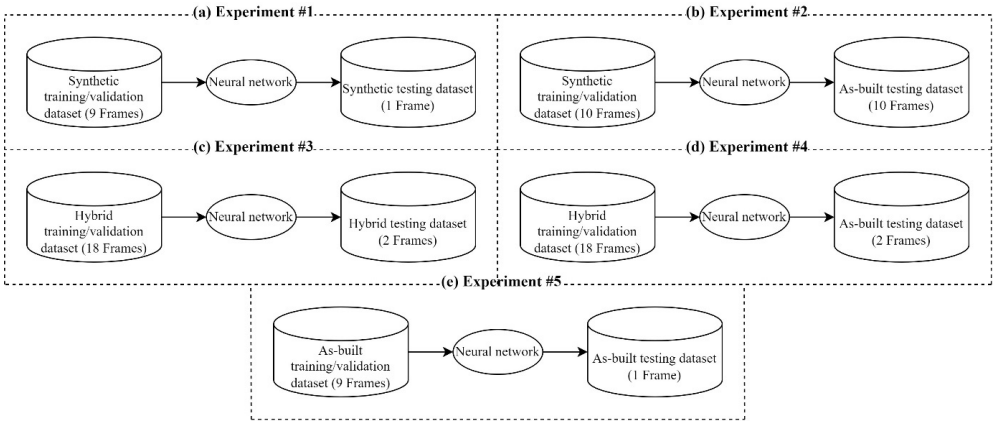


Figure 4. Training, validation, and testing datasets used in each experiment.

- Experiment #1: Using synthetic data for both training/validation and testing ensures that the model's performance is assessed on data similar to what it was trained on, allowing for an evaluation of how well it generalizes within the synthetic dataset's characteristics.
- Experiment #2: Employing synthetic data for training/validation and real as-built data for testing measures the model's ability to generalize from synthetic to real-world scenarios, evaluating its performance on unseen real data.
- Experiment #3: Utilizing a mix of synthetic and as-built data for both training/validation and testing allows for assessing the model's performance within a combined dataset and evaluating its capabilities in both BIM and real-world scenarios.
- Experiment #4: Using a hybrid dataset for training/validation and as-built mixed-reality-captured data for testing assesses the potential of integrating synthetic BIM data to compensate for the limited availability of real-world data, gauging BIM data's applicability in practical scenarios.
- Experiment #5: Training, validation, and testing on as-built mixed reality capture data provide an insight into the model's performance, without any influence from synthetic data.

Class weights, w, are assigned to each label to account for an imbalanced dataset, as shown in Equation (6), whereby P_{total} is the total number of samples, n_{class} is the number of unique classes in the dataset, and f_{class} is the frequency of that class in the dataset.

$$w = \frac{P_{total}}{n_{class} \times f_{class}} \tag{6}$$

As the number of bracing elements is relatively lower than other structural elements in a typical LSF design, this adjustment is necessary to mitigate biases, improving the models' ability to generalize across all classes. The network architecture used in this study is summarized below in Figure 5.

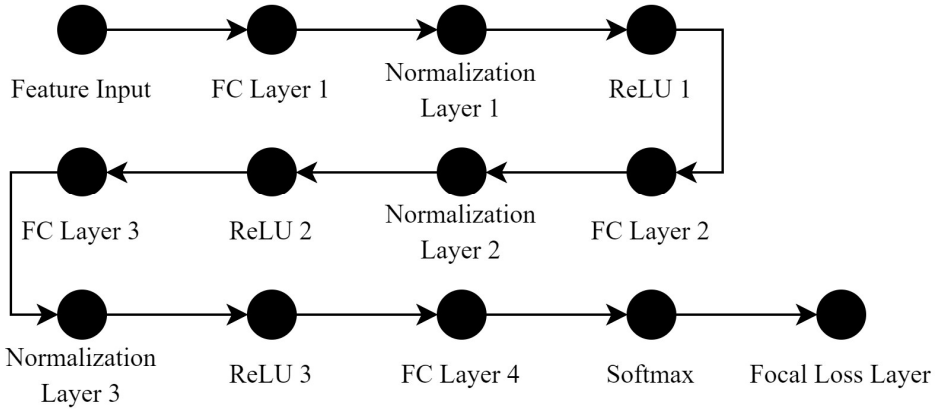


Figure 5. Network architecture.

The first fully connected layer network had 1,024 neurons, followed by 512 and 256 neurons in the subsequent fully connected layers, and their activations were normalized. We also introduced non-linearity by outputting zero for negative input values and passing positive values directly through. A softmax activation function was used before the focal loss layer, whereby the focal loss function, L, was computed using Equation (7) [37].

$$L = -\frac{1}{M}\sum_{m=1}^{M}\sum_{k=1}^{K} T_{mk}\alpha(1 - Y_{mk})^{\gamma}\ln(Y_{mk}) \tag{7}$$

K is the number of classes, whereas M is the number of observations: the number of elements along the first two dimensions of Y multiplied by the number of the anchor box. The ADAM algorithm was used in this study, with an initial learning rate of 0.001 that

decays by a factor of 0.5 every 2 epochs over a total of 20 epochs. A ridge regression, L_2, with a value of 1×10^{-6} was used to add a penalty term to the loss function.

Several metrics were also used to evaluate the model's performance, including accuracy, precision, recall, and $F1$ score, as shown in Equations (8)–(11).

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{Total\ Instances\ (TI)} \tag{8}$$

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \tag{9}$$

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \tag{10}$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

The combination of these metrics is particularly beneficial to gain a comprehensive understanding of the model's strengths and weaknesses.

4. Results and Discussion

In presenting the findings from the experiment outlined in the research methodology, this section aims to draw connections between all the experiments and determine whether synthetic BIM data are useful in real-world applications. Figure 6 shows the overall results from all 5 different experiments as described in the methodology.

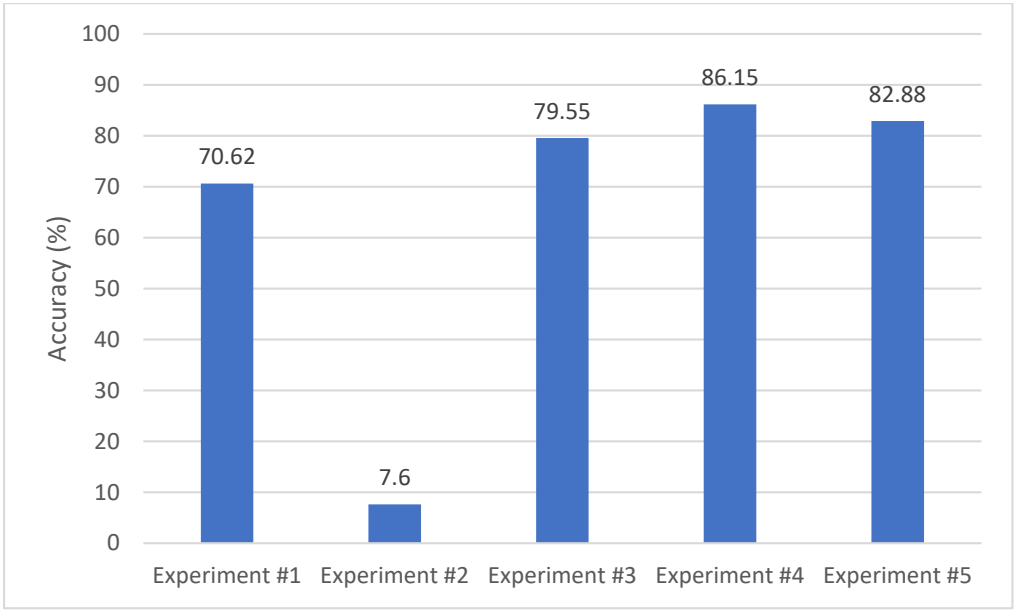


Figure 6. Experimental results.

Using solely synthetic BIM data for the training, validation, and testing of the model in experiment #1, its testing accuracy stood at 70.62%, and this segmented LSF frame can be seen in Figure 7a. This demonstrates the model's ability to learn and adapt to the patterns and features inherent in the synthetic BIM data. The consistency and similarity observed within the synthetic BIM dataset showed little discrepancies or inconsistencies, ensuring that the synthetic BIM dataset remains a reliable resource for training. In addition, it also shows potential in other applications of BIM in the AEC industry, including construction coordination [38], planning [39], and decision making [40], especially in cases where there is a lack of information regarding structural elements.

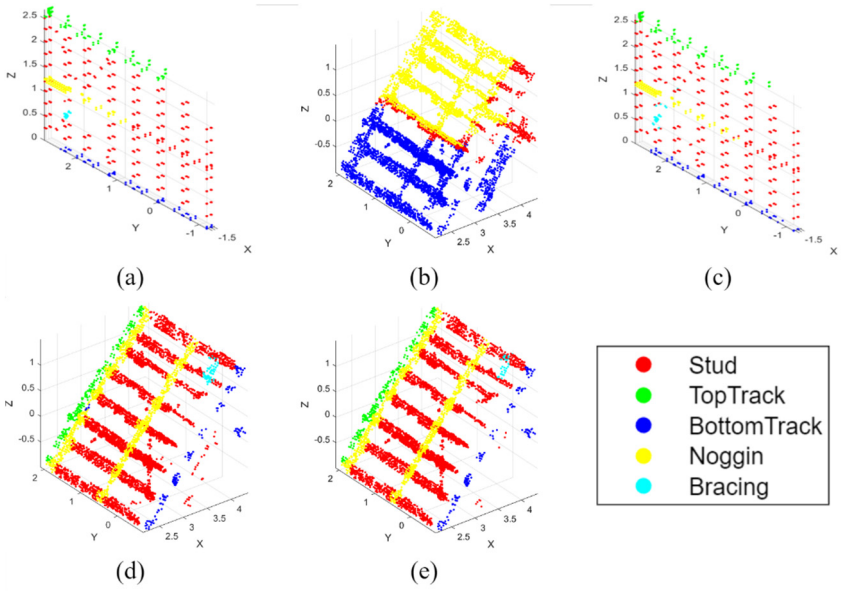


Figure 7. Neural network training results: (**a**) segmentation example from experiment #1; (**b**) segmentation example from experiment #2; (**c**) segmentation example from experiment #3; (**d**) segmentation example from experiment #4; (**e**) segmentation example from experiment #5.

Meanwhile, in experiment #2, we leveraged the synthetic data's advantages of ease of generation and controlled environments to train our model and segment unseen actual as-built mixed-reality-captured data. As the training data were fully synthetic, there is a disparity between the characteristics of the training and testing dataset. As such, the model's tested accuracy averages 7.60%, showing the lack of complexities inherent in the as-built data. An example of its segmented LSF frames can be found in Figure 7b. This clearly showed that the model failed to adapt to real-world applications with solely synthetic data. Nevertheless, this does not imply that synthetic data lack the capability to predict actual as-built data; it is simply that the divergence between both datasets, whether the lack of noise in the synthetic dataset or the differences in density and distribution of the point cloud dataset, led to poor model performance.

In experiment #3, actual as-built data captured from the MR headset were introduced into the training, validation, and testing datasets, forming hybrid datasets. This experiment aims to train a model capable of adapting to the nuanced complexities of as-built data while retaining its segmentation capabilities in a controlled environment. By partitioning the dataset into 16 (80%) LSF frames for training, 2 (10%) for validation, and 2 (10%) for testing, evenly distributed between synthetic and actual frames, we achieved an average testing accuracy of 79.55%. This demonstrates a remarkable model performance on both synthetic and actual data, as seen in Figure 7c. This points towards the model's adaptability in segmenting unseen BIM and MR datasets, with the potential to automate the process of parsing BIM data and the documentation of the existing built environment. It also shows that neural networks contain the capability to automate inspection processes by identifying the discrepancies between as-built and as-planned designs.

In experiment #4, the experiment setup was similar to experiment #3, with the difference being an additional actual as-built frame introduced into the testing dataset. The dataset followed a similar partition, with eight synthetic + eight actual frames for training, two synthetic frames for validation, and two actual frames for testing. The testing accuracy of experiment #4 stood at an average of 86.15%, similar to the previous experiment, and can be seen in Figure 7d. This experiment aims to gauge the usefulness of synthetic data in supplementing the lack of actual data in real life. Therefore, in experiment #5, we trained the neural network solely on actual as-built data to test the applicability of synthetic data. In experiment #5, the dataset was smaller compared to all previous experiments, containing

only 10 LSF frames in total. The resulting testing accuracy of experiment #5 is 82.88%, a slight decrease in model performance compared to experiment #4, which can be seen in Figure 7e. Although the dataset is relatively small, containing 92,319 data points in total, our preliminary results suggest that there is an improvement in model performance when the synthetic dataset is introduced.

Some notable limitations were identified during the data collection process using the MR headset as well. Firstly, the MR headset tends to fail to capture bracing elements, which are 1 mm thick and 30 mm wide, as shown in Figure 8a,b. Secondly, it has difficulty in distinguishing the bottom plate from the floor, as shown in Figure 8b. Its spatial maps often depict rough surfaces, smoothing out any irregularities that might represent thin structural elements, such as bottom plates. Lastly, the raw capture of the MR headset tends to contain floating objects; however, most outliers have been removed using the method described in the research methodology.

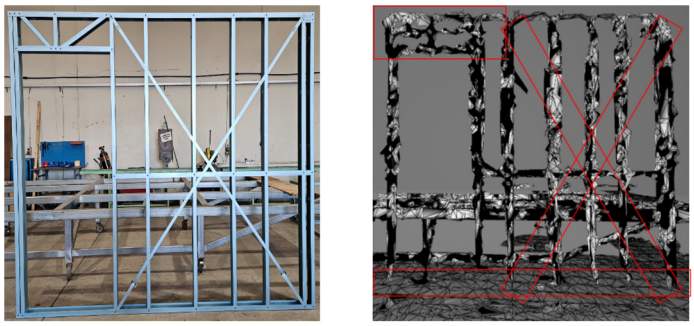


Figure 8. Segmentation limitations: (**a**) actual LSF frame; (**b**) bottom plate and bracing recognition issues with the MR headset.

Nevertheless, the inclusion of synthetic BIM data alongside the actual as-built MR capture data proved to be useful during the model training process, with up to 86.15% accuracy achieved. This indicates the viability of using deep learning techniques with a relatively low cost of data acquisition in the construction monitoring process, relying on BIM data as a source of training datasets. A summary of each experiment can be found in Table 2.

Table 2. Summary of experimental results.

Experiment	Accuracy	Precision	Recall	F1 Score
#1	0.71	0.70	0.65	0.68
#2	0.07	0.06	0.06	0.06
#3	0.79	0.75	0.66	0.70
#4	0.86	0.85	0.75	0.80
#5	0.82	0.76	0.70	0.73

From Table 2, it is clearly seen that the combination of both synthetic and actual datasets aid in model performance. Without relying on synthetic data for model training, the segmentation accuracy decreased from 86% to 82%. However, it should be noted that by relying fully on synthetic data, the model is unable to classify actual as-built LSF frames accurately, as shown in experiment #2.

5. Conclusions

In summary, this study explored the possibility of segmenting the 3D point cloud of an LSF system using both synthetic data obtained from BIM and actual as-built data

obtained directly from MR headsets. The advantage of using the 3D point cloud input from MR headsets, as opposed to existing 2D approaches, lies in its ability to process spatial information more accurately, allowing for depth perception in future automated inspection-related tasks. This study aimed to ascertain the viability of utilizing synthetic data as a supplementary measure due to the scarcity of actual as-built data due to the extensive resources needed to procure LSF frames. A CNN-based approach was used to segment the LSF frames, facilitating the construction monitoring process. Five distinct experiments were conducted using synthetic and actual data and a hybrid mix of both to train and test the model's performance. We assessed the efficacy of synthetic data generated from BIM models in enhancing a neural network's performance in the point-cloud-based segmentation of an LSF system. The experiments showed that introducing synthetic data improves the model's overall performance, with an average testing accuracy of 86.15%, as opposed to solely relying on actual data, which had an average testing accuracy of 82.88%. This demonstrated the benefits of the hybrid effects of synthetic data and as-built scans obtained directly from MR headsets in actual construction scenarios. Nevertheless, the whole dataset, which contains 10 synthetic and 10 actual LSF frames, only consists of 92,319 data points. Therefore, further study should be conducted before drawing a firm conclusion based on these results. In addition, the model trained on a hybrid dataset demonstrates an ability to recognize both BIM data and as-built data captured using the MR headset at an average accuracy of 79.55%. This enables the real-time monitoring of construction progress, allowing for the immediate identification of structural deviations from the planned designs. In future, the authors hope to extend this study to cover a wider range of structural elements, such as roofing systems constructed using LSF. In addition, it would be beneficial to study the applicability of the automated segmentation of structural elements to other structural systems, such as concrete structures. Finally, although the current developed methodology can only recognise these types of structural elements, the authors hope to utilise this point cloud segmentation to develop an automated inspection system to recognise the dimensional and geometrical properties of every structural element in the near future. The authors strongly believe that visual scene understanding using AI will improve the efficiency of progress monitoring and inspection tasks, playing a pivotal role in construction's digitization.

Author Contributions: Conceptualization, Y.S.L., A.R. and A.T.; methodology, Y.S.L. and A.R.; software, Y.S.L.; validation, Y.S.L., A.R., A.T. and D.K.; formal analysis, Y.S.L. and A.T.; investigation, Y.S.L. and A.R.; writing—original draft preparation, Y.S.L.; writing—review and editing, A.R., A.T. and D.K.; supervision, A.R., A.T. and D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the commercial sensitivity of the ENDUROFRAME®® systems.

Acknowledgments: The authors would like to acknowledge ENDUROFRAME®, BlueScope Steel Limited, and Steer Manufacturing for providing 3D and actual light-gauge steel framing system components for this study.

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

References

1. Fang, J.; Li, Y.; Liao, Q.; Ren, Z.; Xie, B. Construction Progress Control And Management Measures Analysis. *Smart Constr. Res.* **2018**, *2*, 1–5. [CrossRef]
2. Pour Rahimian, F.; Seyedzadeh, S.; Oliver, S.; Rodriguez, S.; Dawood, N. On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning. *Autom. Constr.* **2020**, *110*, 103012. [CrossRef]
3. Davila Delgado, J.M.; 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]

4. Kim, J.; Ham, Y.; Chung, Y.; Chi, S. Systematic Camera Placement Framework for Operation-Level Visual Monitoring on Construction Jobsites. *J. Constr. Eng. Manag.* **2019**, *145*, 04019019. [CrossRef]
5. Kim, J.; Hwang, J.; Chi, S.; Seo, J. Towards database-free vision-based monitoring on construction sites: A deep active learning approach. *Autom. Constr.* **2020**, *120*, 103376. [CrossRef]
6. Maharana, K.; Mondal, S.; Nemade, B. A review: Data pre-processing and data augmentation techniques. *Glob. Transit. Proc.* **2022**, *3*, 91–99. [CrossRef]
7. Kim, J.; Kim, D.; Shah, J.; Lee, S. Training a Visual Scene Understanding Model Only with Synthetic Construction Images. In *Computing in Civil Engineering 2021*; ASCE Press: Reston, VI, USA, 2021; pp. 221–229.
8. Guo, M.; Huang, H.; Zhang, W.; Xue, C.; Huang, M. Assessment of RC Frame Capacity Subjected to a Loss of Corner Column. *J. Struct. Eng.* **2022**, *148*, 04022122. [CrossRef]
9. Shi, M.-L.; Lv, L.; Xu, L. A multi-fidelity surrogate model based on extreme support vector regression: Fusing different fidelity data for engineering design. *Eng. Comput.* **2023**, *40*, 473–493. [CrossRef]
10. Milgram, P.; Kishino, F. A taxonomy of mixed reality visual displays. *IEICE Trans. Inf. Syst.* **1994**, *77*, 1321–1329.
11. Feiner, S.; MacIntyre, B.; Höllerer, T.; Webster, A. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Pers. Technol.* **1997**, *1*, 208–217. [CrossRef]
12. Honkamaa, P.; Siltanen, S.; Jäppinen, J.; Woodward, C.; Korkalo, O. Interactive outdoor mobile augmentation using markerless tracking and GPS. In Proceedings of the Virtual Reality International Conference (VRIC), Laval, France, 18–20 April 2007; pp. 285–288.
13. Zaher, M.; Greenwood, D.; Marzouk, M. Mobile augmented reality applications for construction projects. *Constr. Innov.* **2018**, *18*, 152–166. [CrossRef]
14. El Ammari, K.; Hammad, A. Remote interactive collaboration in facilities management using BIM-based mixed reality. *Autom. Constr.* **2019**, *107*, 102940. [CrossRef]
15. Dai, F.; Olorunfemi, A.; Peng, W.; Cao, D.; Luo, X. Can mixed reality enhance safety communication on construction sites? An industry perspective. *Saf. Sci.* **2021**, *133*, 105009. [CrossRef]
16. Lee, Y.S.; Rashidi, A.; Talei, A.; Jian, B.; Rashidi, S. A Comparison Study on the Learning Effectiveness of Construction Training Scenarios in a Virtual Reality Environment. *Virtual Worlds* **2023**, *2*, 36–52. [CrossRef]
17. Waugh, L.; Rausch, B.; Engram, T.; Aziz, F. Inuvik super school VR documentation: Mid-project status. In *Cold Regions Engineering 2012: Sustainable Infrastructure Development in a Changing Cold Environment*; ASCE Press: Reston, VI, USA, 2012; pp. 221–230.
18. Juang, J.; Hung, W.; Kang, S. Kinesthetic and stereoscopic vision for crane training systems. In Proceedings of the 11th International Conference on Construction Applications of Virtual Reality (CONVR) Conference, Weimar, Germany, 3–4 November 2011.
19. Wang, X.; Dunston, P.S.; Skiniewski, M. Mixed reality technology applications in construction equipment operator training. In Proceedings of the 21st International Symposium on Automation and Robotics in Construction (ISARC 2004), Jeju, Republic of Korea, 21–25 September 2004; pp. 21–25.
20. Zhang, J.; Long, Y.; Lv, S.; Xiang, Y. BIM-enabled Modular and Industrialized Construction in China. *Procedia Eng.* **2016**, *145*, 1456–1461. [CrossRef]
21. Li, C.Z.; Xu, X.; Shen, G.Q.; Fan, C.; Li, X.; Hong, J. A model for simulating schedule risks in prefabrication housing production: A case study of six-day cycle assembly activities in Hong Kong. *J. Clean. Prod.* **2018**, *185*, 366–381. [CrossRef]
22. Qi, B.; Razkenari, M.; Costin, A.; Kibert, C.; Fu, M. A systematic review of emerging technologies in industrialized construction. *J. Build. Eng.* **2021**, *39*, 102265. [CrossRef]
23. Lee, Y.S.; Rashidi, A.; Talei, A.; Arashpour, A.P.M.; Pour Rahimian, F. Integration of deep learning and extended reality technologies in construction engineering and management: A mixed review method. *Constr. Innov.* **2022**, *22*, 671–701. [CrossRef]
24. Martinez, P.; Ahmad, R.; Al-Hussein, M. A vision-based system for pre-inspection of steel frame manufacturing. *Autom. Constr.* **2019**, *97*, 151–163. [CrossRef]
25. Wang, L.; Zhao, Z.; Wu, X. A Deep Learning Approach to the Classification of 3D Models under BIM Environment. *Int. J. Control. Autom.* **2016**, *9*, 179–188. [CrossRef]
26. Li, D.; Nie, J.-H.; Wang, H.; Ren, W.-X. Loading condition monitoring of high-strength bolt connections based on physics-guided deep learning of acoustic emission data. *Mech. Syst. Signal Process.* **2024**, *206*, 110908. [CrossRef]
27. 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]
28. Justus, D.; Brennan, J.; Bonner, S.; McGough, A.S. Predicting the computational cost of deep learning models. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 3873–3882.
29. Lashgari, E.; Liang, D.; Maoz, U. Data augmentation for deep-learning-based electroencephalography. *J. Neurosci. Methods* **2020**, *346*, 108885. [CrossRef] [PubMed]
30. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]
31. Alkhalifah, T.; Wang, H.; Ovcharenko, O. MLReal: Bridging the gap between training on synthetic data and real data applications in machine learning. *Artif. Intell. Geosci.* **2022**, *3*, 101–114. [CrossRef]
32. Rajotte, J.-F.; Bergen, R.; Buckeridge, D.L.; El Emam, K.; Ng, R.; Strome, E. Synthetic data as an enabler for machine learning applications in medicine. *iScience* **2022**, *25*, 105331. [CrossRef] [PubMed]

33. Chalhoub, J.; Ayer, S.K. Using Mixed Reality for electrical construction design communication. *Autom. Constr.* **2018**, *86*, 1–10. [CrossRef]
34. Riexinger, G.; Kluth, A.; Olbrich, M.; Braun, J.-D.; Bauernhansl, T. Mixed Reality for On-Site Self-Instruction and Self-Inspection with Building Information Models. *Procedia CIRP* **2018**, *72*, 1124–1129. [CrossRef]
35. Ren, B.; Wang, H.; Wang, D.; Guan, T.; Zheng, X. Vision method based on deep learning for detecting concrete vibration quality. *Case Stud. Constr. Mater.* **2023**, *18*, e02132. [CrossRef]
36. Truong, V.D.; Xia, J.; Jeong, Y.; Yoon, J. An automatic machine vision-based algorithm for inspection of hardwood flooring defects during manufacturing. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106268. [CrossRef]
37. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2999–3007.
38. Singh, M.M.; Sawhney, A.; Borrmann, A. Modular coordination and BIM: Development of rule based smart building components. *Procedia Eng.* **2015**, *123*, 519–527. [CrossRef]
39. Hammad, A.W.; Akbarnezhad, A.; Rey, D.; Waller, S.T. A computational method for estimating travel frequencies in site layout planning. *J. Constr. Eng. Manag.* **2016**, *142*, 04015102. [CrossRef]
40. Chai, C.; Tan, C.; Aminudin, E.; Loo, S.; Goh, K.; Theong, M.; Lee, X.; Chin, L. The potential cost implications and benefits from Building Information Modeling in Malaysian construction industry. In Proceedings of the ASIA International Multidisciplinary Conference, Johor Bharu, Malaysia, 1–2 May 2017 .

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.