buildings

Article

Feature-Based Deep Learning Classification for Pipeline Component Extraction from 3D Point Clouds

Zhao Xu [1,*], Rui Kang [1] and Heng Li [2]

[1] Department of Civil Engineering, Southeast University, Nanjing 210096, China; kr97102@seu.edu.cn
[2] Department of Building and Real Estate, Hong Kong Polytechnic University, Kowloon, Hong Kong 999077, China; heng.li@polyu.edu.hk
* Correspondence: xuzhao@seu.edu.cn

Abstract: This paper proposes a novel method for construction component classification by designing a feature-based deep learning network to tackle the automation problem in construction digitization. Although scholars have proposed a variety of ways to achieve the use of deep learning to classify point clouds, there are few practical engineering applications in the construction industry. However, in the process of building digitization, the level of manual participation has significantly reduced the efficiency of digitization and increased the application restrictions. To address this problem, we propose a robust classification method using deep learning networks, which is combined with traditional shape features for the point cloud of construction components. The proposed method starts with local and global feature extraction, where global features processed by the neural network and the traditional shape features are processed separately. Then, we generate a feature map and perform deep convolution to achieve feature fusion. Finally, experiments are designed to prove the efficiency of the proposed method based on the construction dataset we establish. This paper fills in the lack of deep learning applications of point clouds in construction component classification. Additionally, this paper provides a feasible solution to improve the construction digitization efficiency and provides an available dataset for future work.

Keywords: deep learning; pipeline component extraction; point clouds; feature; CNN (convolutional neural network)

check for updates

Citation: Xu, Z.; Kang, R.; Li, H. Feature-Based Deep Learning Classification for Pipeline Component Extraction from 3D Point Clouds. *Buildings* **2022**, *12*, 968. https://doi.org/10.3390/buildings12070968

Academic Editor: Ahmed Senouci

Received: 14 June 2022
Accepted: 5 July 2022
Published: 7 July 2022

1. Introduction

The 3D reconstruction models are gradually replacing 2D drawings for information transmission and more in-depth processing to meet the demands of civil engineering digitization, according to the research by Ma and Liu [1]. Among a variety of 3D information data formats, the point cloud model is the research focus of many scholars. The 3D point cloud data are widely used in the construction industry for model reconstruction, geometry inspections and other applications, but there is still a research gap regarding the practical applications [2]. In the early stage of point cloud data processing, scholars mainly utilize traditional algorithms to deal with complex and irregular object reconstruction [3], complicated scenes with repetitive objects [4] and the updating of as-designed BIM to as-built BIM [5]. On the basis of shape extraction, some scholars further enriched the semantics of point cloud data and deepened the relationship between BIM [6,7] and point clouds via IFC (Industry Foundation Classes) extension [8]. This has had positive significance for data management in the field of civil engineering, and has greatly promoted the informatization process in the construction industry. However, such studies mainly focus on indoor scenes and generally require manual participation, which leads to a low level of automation and compatibility [9]. Thus, some scholars have turned to deep learning for resolution. With the maturity of deep learning algorithms, more network structure designs for the deep learning of point clouds have emerged, the feasibility of applying deep networks to point cloud learning has been verified and the design theory has been continuously improved [10]. The application of deep learning in civil engineering information management has developed from image damage detection [11] to 3D reconstruction segmentation [12]. In order to use deep learning to enhance the efficiency of engineering information acquisition, some scholars directly apply deep learning networks on IFC models [13] or use deep learning to assist in removing unwanted information [14]. Overall, most scholars focus their attention on deep learning to help in the acquisition and processing of 3D point clouds, so as to reduce the manual participation and achieve automation. This approach would be for the great promotion of efficient digital management, especially for complicated building scenarios such as MEP (Mechanical, Electrical and Plumbing) systems.

With the development of 3D reconstruction technology, the accuracy of the obtained data is gradually improved. At this time, the requirements for the accuracy and flexibility of the model become more stringent. In the realization of as-built BIM from point clouds, the extraction of building components plays a vital role, among which pipeline component extraction is an important task. The digitization of existing buildings can be completed by extracting and identifying point clouds, and then the management of important building components, such as MEP systems, can be completed. In the construction industry, the application of point cloud data in information sharing platforms is synchronized with the development of point cloud data processing methods [15]. To efficiently use point clouds, scholars have paid attention to semantic recognition, which is an essential step to identify the part that a point cloud belongs to. In reverse engineering, many algorithms are used in the mesh and point clouds for labeling, such as randomized cuts for the mesh [16], mesh labeling via CRF [17] and the octree-based method for point cloud segmentation [18]. These algorithms are designed using the geometry features and can work for certain datasets, but it is difficult for them to maintain high accuracy in various datasets. To address this problem, the concepts and structures of deep learning are introduced to design a self-learning algorithm for mesh labeling [19], and we further turn to point cloud labeling [20] due to the flexibility and integrity of the point cloud approach.

However, the currently proposed algorithms experiment with repetitive and limited datasets, which leads to difficulties for actual applications in engineering practice, with even fewer applications in MEP systems [9]. Among the datasets used by these algorithms, one part of the datasets is composed of small-volume point clouds generated by CAD models which have little noise interference, such as ModelNet40 [21], while the other part of the datasets is designed for scene segmentation and focuses on identifying ceilings, tables and chairs in indoor rooms, such as the Stanford 3D semantic parsing dataset [22]; or trees, roads and buildings in outdoor spaces, such as Semantic3D.net [23]. In addition to the problems with the datasets used for training, the practical engineering application of these algorithms is also restricted by the environment and other conditions. The point clouds obtained on site are incomplete because of interference. At the same time, they are affected by the speed of the data collection, meaning it is difficult to obtain complete attribute data, such as RGB (Red, Green, Blue) data.

To overcome the above problems, it is necessary to adjust the structure design of the deep learning algorithms to adapt to the needs of engineering applications, especially for MEP systems. Thus, based on previous studies, this paper proposes a new deep network structure and builds a dataset that emphasizes engineering scenarios for learning and training. The key to our approach is the surrounding description of a single point. Through the usage of the SHOT (signature of histograms of orientations) and spin image approaches, the attributes of a single point are expanded to fill the gaps in data collection.

The key contributions of our work are as follows: (1) we design a novel neural network architecture suitable for an imperfect point cloud collected from the construction project; (2) we introduce the concepts of SHOT and spin image in point cloud deep learning and improve the performance of point cloud labeling by depicting the distribution of points; (3) we establish a point cloud deep learning dataset for engineering application scenarios and pipeline component extraction.

The rest of this paper is organized as follows. Section 2 reviews related work and Section 3 introduces the overall methodology used for the point cloud labeling. In Section 4, the architecture of the network is shown to express how it works. Then, we demonstrate the performance in the experiments in Section 5 with pipelines as examples. Finally, the article is concluded in Section 6.

2. Literature Review

2.1. Information Extraction and 3D Reconstruction of Pipeline Components

As a more efficient data acquisition method, non-destructive tests (NDTs) replace traditional manual measurement methods in many application scenarios [24]. Among the many NDT methods, oblique photography, ultrasonic detection and laser scanning are the most representative ones. As a large-scale data collection method, oblique photography mainly uses drones for outdoor data collection. The subsequent data processing tends to distinguish buildings from the environment [25] or extract building outlines [26]. Ultrasonic detection is mainly used for internal damage detection, such as crack detection [27] and weld detection [28], and no reconstruction model is generated during the detection process. Compared with the first two types of NDTs, laser scanning is more suitable for construction component reconstruction in the field of civil engineering due to its accuracy and scanning range. Different types of laser scanning equipment can be applied for overall scene scanning or partial scanning to obtain stable and accurate point clouds. Using these point clouds, 3D models of buildings or components can be reconstructed to conduct measurement [29], recognition [30] and other data processing operations.

Processed point clouds are used in the field of civil engineering for quality and safety management during the construction phase and in building information management, which is known as BIM. Containing high-precision three-dimensional data, the point cloud can describe the settlement state of the construction site and the deformation state of the construction components, which is of great benefit for the collection of data for MEP systems from complicate pipelines. Some scholars have proposed algorithms, including the convex hull [31] and improved RANSAC algorithms [32], to extract primitive geometric information form a construction site or building for further as-built BIM applications.

As an important building information exchange technology, BIM requires more data to show a complete model with rich semantics to support the information required from design to operation and maintenance phases while gradually improving. With the expansion of the application range and the increase in importance, there are stricter requirements for the point cloud data used in the BIM platform [33]. In the process of using BIM for digital construction, pipeline management methods have also been developed, since the pipeline occupies an important position during maintenance. In buildings that use BIM technology for forward design, the management of pipelines can achieve efficient and diversified purposes. In addition to conventional pipeline information management and collision detection, some scholars have realized the prediction of pipeline corrosion on the basis of BIM. Tsai et al. [34] utilized the semantic information management function to store pipeline sensor information and to monitor and visualize the maintenance status. Under the premise of complete pipeline information in BIM, many technologies, e.g., IoT and RFID, can be used to assist in pipeline management [35,36]. Although BIM technology has been widely used in the construction industry, a large number of existing buildings need to be processed via as-built BIM technology to harness its superior functions for efficient management. For the realization of as-built BIM for pipelines, there has been continuous research for the purpose of reducing the manual participation proportion during processing. Patil et al. [37] paid attention to the Hough transform for the automatic detection of cylinder parameters in point clouds. Tran et al. [38] used traditional shape features for cylinder fitting to complete the pipeline extraction in a similar way.

Among these works for MEP systems and pipeline management, these proposed algorithms are usually used in ideal conditions, while the point clouds collected in actual engineering scenes, especially in complex MEP systems, are often noisy and incomplete.

The existing research has explored a feasible workflow to use the point clouds for building information acquisition and management. However, these traditional point cloud processing methods require a lot of manual participation and are limited to a few fixed models, leading to poor efficiency and compatibility when dealing with the pipelines.

2.2. Deep Learning in Point Cloud

As a data format that is widely used in 3D reconstruction, the point cloud has always been a research hotspot in graphics processing and computer vision. Recognizing the semantic information of the point cloud as an important part of its use has been a continuous concern for scholars [39]. Before deep learning gained widespread attention, graphical analyses of point clouds were already carried out. The point cloud data features, including spin image [40] and SHOT [41] features, are calculated to classify different parts to achieve segmentation. Traditional classification algorithms such as the principal component analysis [42] are also used to assist in point cloud processing. Although these characteristics can help distinguish the geometric features of a point cloud reconstruction model, they are too rigid to lose their self-learning ability and strong adaptability. The emergence of deep learning algorithms fills the gaps in this research field.

Deep learning for point clouds mainly focuses on the research fields of recognition and classification and is applied to road condition recognition for automatic driving [43], large-scale scene object classification [44] and indoor scene recognition [23]. According to the methods of deep learning data recognition, they can be divided into three types in general: Pointnet, Voxelnet and feature-based net.

Pointnet proposed by Qi et al. is an important algorithm in the field of point cloud deep learning [20]. Pointnet is highly efficient and robust and can deal with object classification, part segmentation and scene semantic parsing. Since Qi et al. provided both an authoritative theoretical analysis and experimental evaluation, many follow-up works aiming to improve the network structure have been based on Pointnet, such as Pointnet++ [10], Foldingnet [45] and dynamic graph CNN [46]. The novel deep net architecture proposed as Pointnet, which focuses on processing unordered 3D points, has proved its stability and efficiency via experiments. Compared with other algorithms, this algorithm is characterized by the extraction and processing of the main properties of a point cloud, which are the disorder, interactions among points and invariance under transformations. The core network structures of Pointnet are T-net, which is designed to strengthen the relevance of points in the early stage of data training, and the max pool layer, which is used for dealing with unordered points and keeping the invariance under transformations. Additionally, on the basis of a classification network, Qi et al. [20] further expanded the network structure to implement part and semantic segmentation by combining point features and global features. Intuitive explanations were developed for the robust and effective performance of Pointnet in their paper, and more studies are following the steps of Pointnet to pursue better performance and more flexible use in more fields.

However, because of the training samples used by Pointnet, which are relatively simple and small in size, and the lack of attention to local features, it is difficult to implement Pointnet in large-scale continuous scenes, especially construction engineering scenes.

Differing from Pointnet, which uses a small point cloud as the input, Voxelnet emphasize points connections. This algorithm is characterized in the early-stage data by voxel mark processing. Point clouds are allocated in voxels and form input data according to the distribution in the voxel [47] or voxel labels [48]. Voxelnet can achieve precise control because of the rasterization of point clouds and can handle variable point clouds in more flexible ways in different usage scenarios. Although forming a voxel grid in the most straight-forward way fully utilizes ConvNet, which was originally designed for 2D imaging by changing the input data, it also leads to many disadvantages. Since VoxelNet is mainly used in large-scale scenarios, it is often difficult to balance the workload and algorithm accuracy, and the stability is greatly affected by the voxels. Although Kd-net was proposed to further optimize the low efficiency of the voxels, it is still not a fundamental solution to the application of deep learning on 3D data based on the characteristics of the point cloud [49].

In addition to the above two types of deep learning networks, some scholars start from combining deep learning and traditional shape features by extracting features in advance and then using them as the inputs of networks [19,50]. These types of networks rely on the study foundation of traditional algorithms to provide reliable parameters for subsequent learning, balancing the robustness and flexibility. Nevertheless, there have been cases where traditional shape features are overused. For example, Guo et al. used more than 7 features in their paper, which led to excessive reliance on the traditional shape features and weakness of the learning network, together with increasing the workload at the preprocessing stage. Although these networks can provide stable performance, they have not brought about obvious improvements compared to traditional algorithms.

2.3. Deep Learning in Construction Industry

The introduction of digital technology such as deep learning has accelerated the construction digitization process and provided more efficient tools to process engineering data. A recent study [51] showed that deep learning technologies have been applied for prevalent construction challenges such as site management, budget and energy control and building quality monitoring. In terms of information management and prediction, Sun et al. [52] used the long–short-term memory (LSTM) neural network in deep learning for exchange rate forecasting and Ziari et al. [53] made full use of the deep learning network of natural language processing (NLP) to assist highway agencies in decision-making procedures, indicating that deep learning is gradually shifting from theory to practice and becoming a substitute for traditional algorithms in some empirical research fields. Additionally, during the process of image recognition, since ImageNet [54], which is used for large-scale image recognition, was proposed in 2012, deep-learning-related research has received more attention and there has been wide adoption of similar deep learning structures, e.g., convolutional neural networks. In the construction industry, the numerical prediction and image recognition functions of deep learning algorithms are mainly used to solve practical engineering problems. For instance, Deng et al. [55] and Nguyen et al. [56] separately proposed CNN model and DNN model to predict the strength of concrete for monitoring and design. Rahman et al. [57] designed an RNN (recurrent neural network) model for building energy prediction. Rafiei and Adeli [58] presented a novel machine learning model for estimating construction costs by combining the DBM-SoftMax (deep Boltzmann machine) layer and BPNN (back-propagation neural network). For image recognition functions, scholars tend to invest more in research on on-site worker behavior and construction quality and safety monitoring. For example, Fang et al. [59,60] proposed a CNN model to check workers' posture and helmet-wearing from videos. Kolar et al. [61] presented a CNN-based model to detect guardrails in 2D images for safety improvements. In researching applications related to quality issues, scholars have focused more on crack detection. Similar CNN-based approaches have been adopted for crack detection, where scholars have made efforts towards accuracy and robustness improvements [62,63].

Although scholars have done a lot of work by combining deep learning approaches for the digitalization of the construction industry, there has been very little work related to 3D scenes and BIM, despite the model recognition work carried out by Wang et al. in a BIM environment [64]. The adoption of BIM models has brought significant improvements to construction digitization over the years [65]; scholars are also striving to achieve BIM through the use of more advanced equipment and algorithms [66] but there is still a lot of work to be done to achieve efficiency improvements, especially for as-built BIM. At present, the achievement of as-built BIM requires a lot of human participation to manually label construction components, which can be automated by applying deep learning algorithms for point clouds. However, the datasets used in the related algorithms, such as the ShapeNet dataset [67], ModelNet40 dataset [21] and Stanford large-scale indoor space dataset [22], do not match the engineering application scenarios.

3. Methodology

In this paper, we present a point cloud classification method by using deep CNNs, as shown in Figure 1. First, the original data are processed by the preprocessing network to obtain the training features. The original input data are the global coordinate values of the point cloud. Inspired by PointNet, the network structure of this paper also adopts a combination of global features and local features for training, whereby the global features are formed by the global coordinates of a single point, which are entered into UnitNet, then according to the traditional algorithm the spin image, SHOT and normal of each point are calculated as local features and input into FeatureNet for training. Differing from PointNet's processing of global and local features, the different kinds of features in this paper are calculated separately in the early stage and trained by independent networks, which increases the irrelevance of the data. After the two kinds of features are trained on their own networks, 36-dimensional and 64-dimensional data are output, from which the new input data are merged as 100-dimensional data. Second, deep CNNs are built for feature extraction and final classification. Before being sent to FinalNet for training, a 2D feature matrix (n $\times$ 10 $\times$ 10) is transformed from processed features in the form of feature maps to facilitate subsequent feature extraction. Finally, the deep convolutional networks and the pooling layers are combined for training and learning, and two fully connected layers are applied in the end to obtain the classification results.

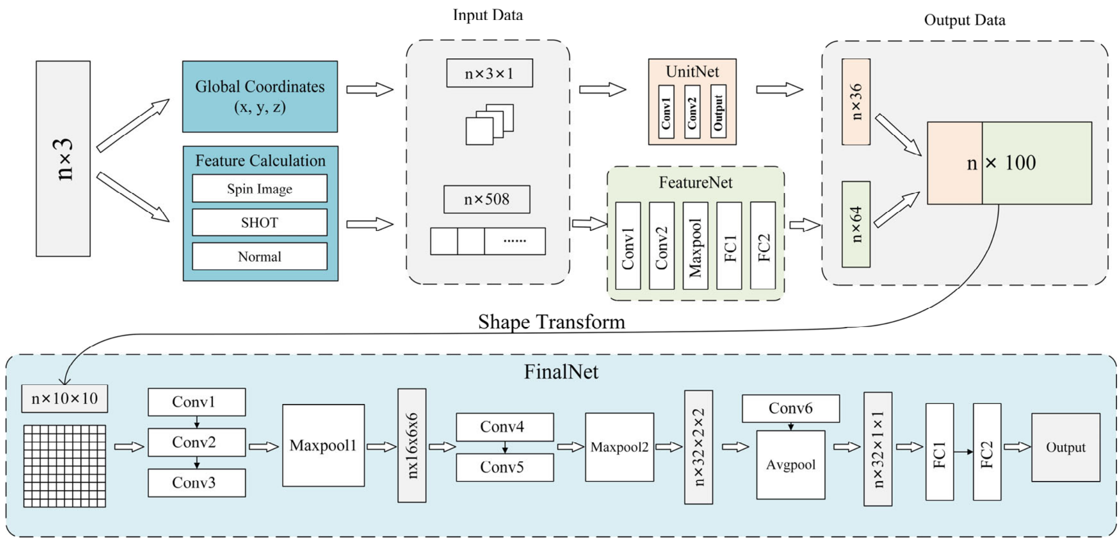

Figure 1. Network resolution.

The key parts of this paper are the preprocessing networks used for the extraction of local and global features by Unitnet and FeatureNet in the early stage, and the use of a deep convolutional network in FinalNet to weight the merged features for recognition in the later stage.

4. Network Architecture

4.1. Preprocessing Networks

In this paper, the network is used to classify the large-scene point clouds, especially the one that is collected at the construction site. Due to the different application scenarios, the data input to the neural network have the following characteristics: (1) Isolation of the input data. The point cloud is disordered, and in the recognition of large scenes, each point is input into the network separately, which highlights the weak connection and poor continuity between the data. (2) The particularity of the application scenario. Many segmentation and recognition algorithms for large scenes are based on indoor scenes, and many elements such as chairs can be obtained from existing model libraries and trained in advance. For a construction site, there are few large-scale point cloud datapoints available for use. (3) The lack of data features. Although the current laser scanning technology has reached a high level of accuracy and good RGB color rendering performance by taking photos for scanning, in order to ensure versatility and avoid interference during color collection under construction conditions, the original data obtained often only have coordinate values.

In order to solve the above problems, this paper sets up UnitNet and FeatureNet for data preprocessing for global features and local features.

4.1.1. Global Feature Extraction

UnitNet is designed for global coordinate transformation, where each point is input as a separate unit with normalized coordinate values as input channels. The network architecture is shown in Figure 2. The input unit is trained by a one-kernel convolutional network with 16 channels and 36 channels. Simultaneously, the BN (batch normalization) layer and ReLU layer are applied for parameter correction and activation during the training process.

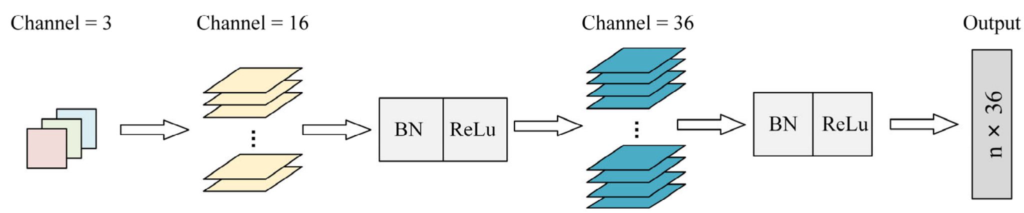

Figure 2. Architecture of UnitNet.

It can be seen from the network structure that the shape of the input unit is not changed, only the number of channels is gradually increased. In this way, the coordinate values of the input data can be further input into the subsequent network after transformation, while retaining the global characteristics of the original data. In this process, the coordinate function of the point is strengthened.

4.1.2. Local Feature Extraction

The raw data from the point cloud have strong independence, which makes it difficult to find the connections with the surrounding points when only using the coordinate value. This disadvantage can be overcome by traditional algorithms such as SHOT and spin image algorithms through the description of local points.

In this paper, SHOT (signature of histograms of orientations) and spin image are important features used as input data to train neural networks, which are originally designed for surface matching. Both features can be used to describe the distribution of surrounding points, whereby SHOT focuses on the locations of surrounding points and spin image focuses on the distribution density of the points.

There are two reasons for using SHOT and spin image as input data for this convolutional neural network: (1) Each point in the point cloud exists independently, and its own attributes need to be obtained through subsequent processing, except for the coordinates and RGB color that are obtained. SHOT and spin image can increase the correlation between points and attach local attributes to independent single points. (2) The CNN framework requires rich data for training and judgment, so algorithms that form a large amount of features are needed.

SHOT

SHOT (signature of histograms of orientations) was proposed as a local reference system for surface matching [41]. SHOT, which is also as a local 3D descriptor that has been used in many scenarios, balances the signature and histogram to maintain the descriptiveness as well as the robustness. After the concept of SHOT was put forward, some scholars further improved it with global structure frames [68] or textures [69]. In this paper, a local reference frame is first established to determine the locations of neighbor points, as shown in Figure 3. In the local reference frame, the spherical space is divided into 32 spatial bins, resulting from 8 latitude divisions (only 4 are shown in Figure 3), 2 longitude divisions and 2 radius divisions. According to the normal of the point in every spatial bin, n_{vi}, and the normal of the feature point, n_p, the cosine of the corresponding angle, θ_i, is calculated by $\cos\theta_i = n_p – n_{vi}$. Then, according to the calculated cosine value, 11-level histogram statistics are performed on the number of points falling into each spatial bin. Finally, after the data are normalized, a 352-dimensional $(8 \times 2 \times 2 \times 11)$ feature is generated as an array to input to the neural network.

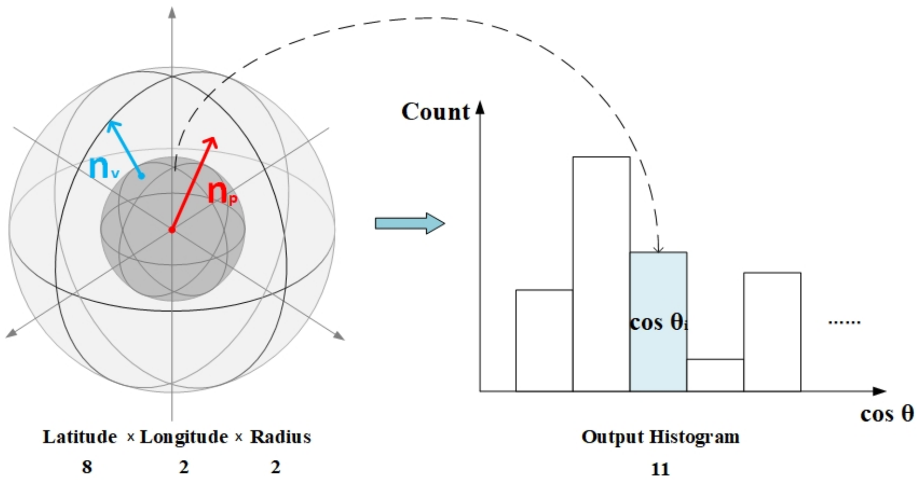

Figure 3. Generation principle of SHOT.

Spin Image

Spin image was proposed by Johnson et al. in 1999 for efficient object recognition and surface matching in 3D scenes. [40] It was first used in 3D meshes. Subsequently, the related algorithm further optimized the spin image approach [70,71]. With the rise of laser scanning technology, point clouds are gradually displayed as 3D models in more application scenarios. Thus, there are increasing studies exploring the application of the spin image approach in point clouds, such as for classification [72] and registration [73]. In addition to the traditional application technology route, a spin image also provides usable data features for deep learning in point cloud processing [19].

The spin image approach was proposed to perform surface matching by depicting the local distribution of other points around the feature point. As shown in Figure 4, first an oriented point with the surface normal is selected to define a coordinate system, where α is the radial coordinate and β is the elevation coordinate. Then, a 2D accumulator indexed by α and β is rotated, taking the normal as the axis. As the accumulator rotates, the number of points falling into the bin indexed by (α, β) gradually increases, forming the spin image, which can be described by Equation (1), where $S_O(x) : R^3 \rightarrow R^2$, p is the reference point, n is normal vector of p and x is the neighbor point.

$$S_O(x) \rightarrow (\alpha, \beta) = \left(\sqrt{\|x-p\|^2 - (n\cdot(x-p))^2}, n\cdot(x-p) \right) \quad (1)$$

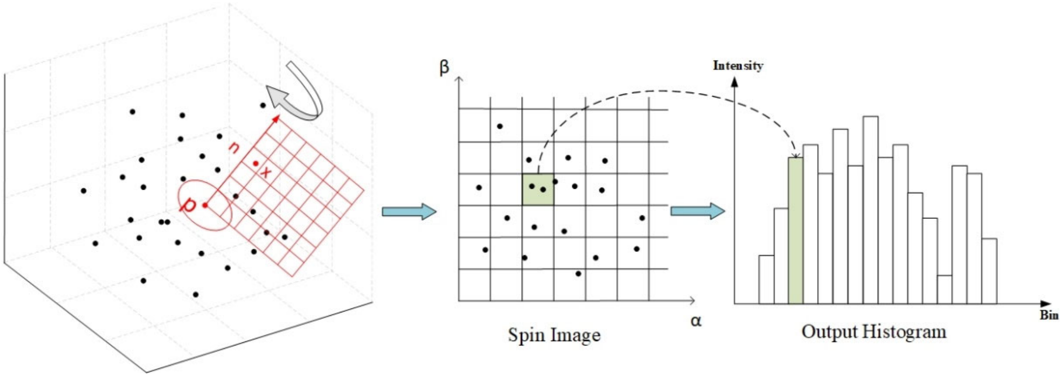

Figure 4. Generation principle of the spin image approach.

Finally, an intensity-related histogram is output as an array to compose the input data of the neural network.

FeatureNet

The output of the above algorithms forms a sequence of 508 elements, of which 352 are from SHOT, 153 are from the spin image and 3 are from the normal. Most of the elements are obtained by establishing a local spatial coordinate system and describing the surrounding points in blocks, which also leads to two problems: (1) Excessive data volume. The method of dividing the space into blocks and the output form of the histogram results in a large amount of local feature data, which is difficult to process. (2) The existence of invalid data. Since the points are not evenly distributed in all spaces, there are lots of zero-value data in the sequence, together with invalid data caused by noise.

The large data volume is often processed using a PCA algorithm for dimensionality reduction, as shown in other studies, but this issue cannot be solved in this paper due to the linearly independence of the feature data. Therefore, this paper proposes the use of FeatureNet for effective feature extraction to reduce the amount of data and eliminate invalid data. The structure of FeatureNet is shown in Figure 5.

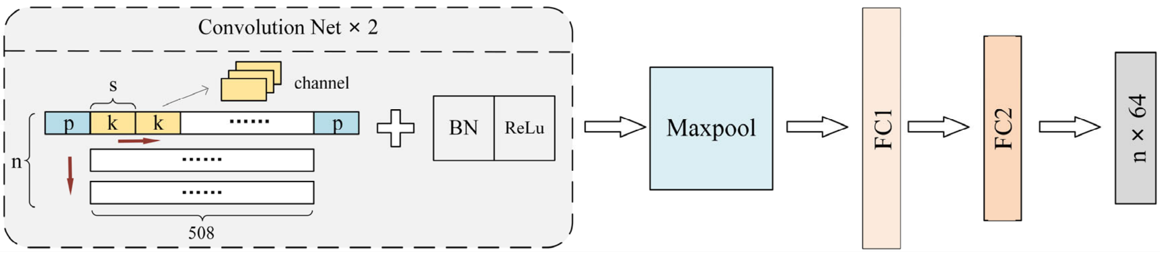

Figure 5. The FeatureNet structure proposed in this study.

In FeatureNet, two one-dimensional convolutional layers and a pooling layer are used to assign weights to each feature element to eliminate invalid data and zero-valued data, and then two fully connected layers are applied to perform dimensional compression to achieve the purpose of data cleaning. Through the processing of FeatureNet, 64-dimensional data are finally output, effectively compressing the dimensions of feature data.

4.2. FinalNet

FinalNet is the main structure that implements feature extraction in the deep convolutional networks. As shown in Figure 1, considering the deeper network structure leads to better performance [74] and improves the response speed of the network structure. FinalNet uses 6 convolutional layers to perform the feature extraction, 3 pooling layers to reduce the dimensions and 2 fully connected layers to make the final judgement.

FinalNet directly processes the feature map size formed by reshaping the sequence spliced from the local feature and global feature. Figure 6 shows the details for composing these two features. There is a weak connection between the two different features in the initial feature map formed via concatenation and deformation in the first stage. After processing the convolutional layer, connections between different feature elements are established through convolutional kernels.

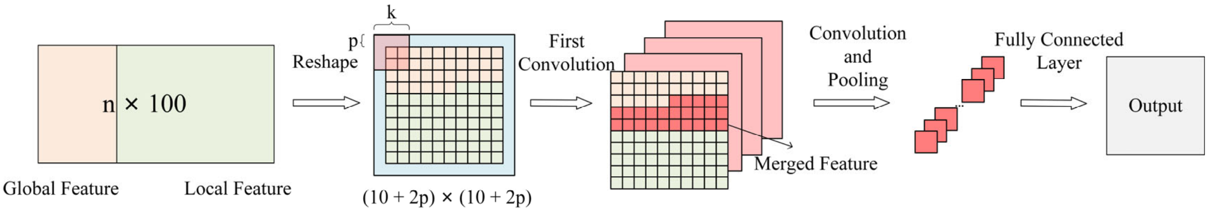

Figure 6. Feature combination process in FinalNet.

The merging result after the first convolution phase can be seen in Figure 6. Subsequently, with the deepening of the convolutional layers and the participation of the pooling layers, a fully integrated feature sequence is formed. Finally, the fully connected layers are applied to form non-linear combination and output the result.

5. Experiment

This paper focuses on the object classification of construction components, especially the components that require maintenance management in complex construction scenes. The existing datasets, such as ModelNet40 (Wu et al., 2015), which has 12,311 CAD models from 40 man-made object categories, and ShapeNet (Yi et al., 2016), which is normally used for indoor scenes in segmented learning, e.g., for tables and chairs, cannot meet the need for the classification of construction components in the field of civil engineering. Additionally, the point clouds generated by CAD models cannot reflect the actual state of the point clouds obtained via laser scanning due to the environmental interference and blind spots. Thus, this paper uses a self-built dataset via the laser scanning and labeling of the appropriate scenes, since there are few datasets that can be directly used for the related training and learning processes.

In this paper, the classification of pipelines, which are difficult to manage in the construction process and complicate the maintenance process, is shown as an example to prove the validity of the proposed network. The whole scene with the format of the point cloud and the pipelines that are manually extracted, labeled and used for training and learning can be seen in Figure 7.

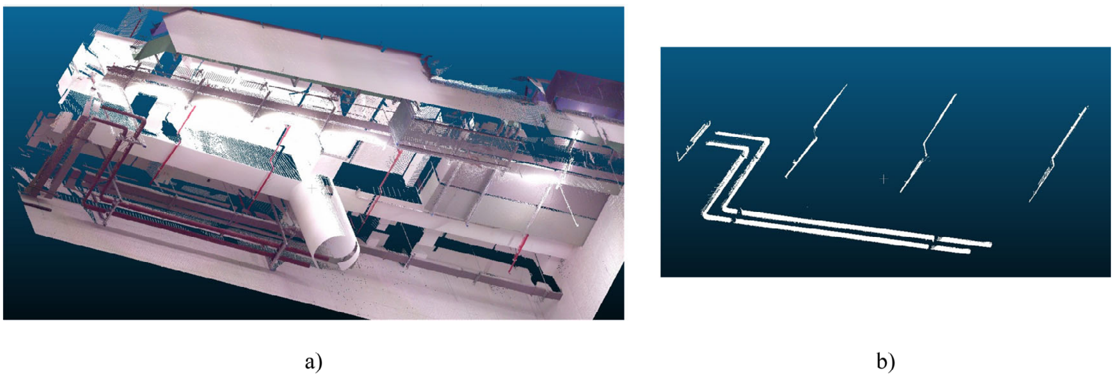

Figure 7. The point cloud dataset. (**a**) is the engineering scene dataset and (**b**) is showing extracted pipelines.

5.1. Architecture Design Validation

In this section, a control experiment is designed for the validation of the networks this paper has proposed. The networks that are trained and perform the pipeline classification in the control experiment are inspired by feature-based DNNs (Y. Fang et al., 2015; Guo et al., 2015), where traditional features are extracted and sent into deep networks directly. In some simple application scenarios (such as image recognition), a network that has enough neurons can achieve good results, even without an effective structural design. However, compared with ordered and feature-rich datasets, such as image datasets, the raw data from point clouds, which are unordered, lack features and are weakly connected with each other, meaning they have higher requirements. Thus, a control experiment is designed here for the validation of UnitNet and FeatureNet, whereby spin image and SHOT features are combined and sent to a deep convolution network directly after the PCA for dimensionality reduction, as shown in Figure 8. This control experiment used for comparison contains 11 convolutional layers, together with BN, ReLU and 2 pooling layers.

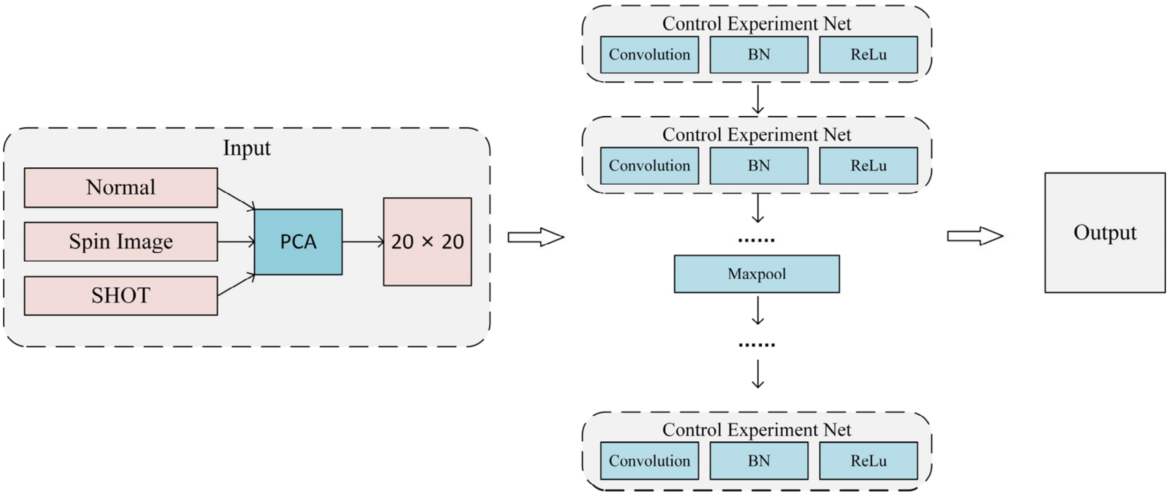

Figure 8. Network structure of the control experiment.

The results of the control experiment demonstrate the positive effects of the application of UnitNet and FeatureNet, as listed in Table 1. The train loss and train accuracy of the control experiment are better than the proposed method but lead to an overfitting result, as reflected by the test accuracy comparison.

Table 1. The results of the pipeline classification experiments.

	Train Loss	**Train Accuracy**	**Test Accuracy**
Control Experiment	0.0036	99.87%	82.37%
Proposed method	0.0054	99.75%	94.62%

5.2. Results of the Dataset and Comparison

Due to the dataset used and the specificity in the engineering of the classification object, there are few experimental comparison datasets available for other algorithms. The experimental results are mainly presented via a large-scale scene and a comparison with the control experiment. The results listed in Table 2 show that our method, which uses UnitNet and FeatureNet for structural improvement, significantly outperforms the deep network in the control experiment.

Table 2. Classification results for a large-scale dataset.

Dataset	Mean	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9
Points number		191,404	696,406	1,419,606	718,361	411,160	296,335	521,597	1,203,240	732,260
Control Experiment	84.06%	80.29%	88.07%	91.55%	92.73%	82.98%	86.10%	92.78%	78.59%	63.45%
Proposed method	98.03%	97.24%	97.50%	94.93%	98.91%	97.80%	98.17%	99.59%	98.54%	99.63%

The qualitative visible results are shown in Figure 9. Here, a comparison of the classification results of the two algorithms is shown for datasets 2 and 5. It can be seen more intuitively that our method can effectively avoid misjudgments by strengthening the connections between local features and global features to achieve better classification.

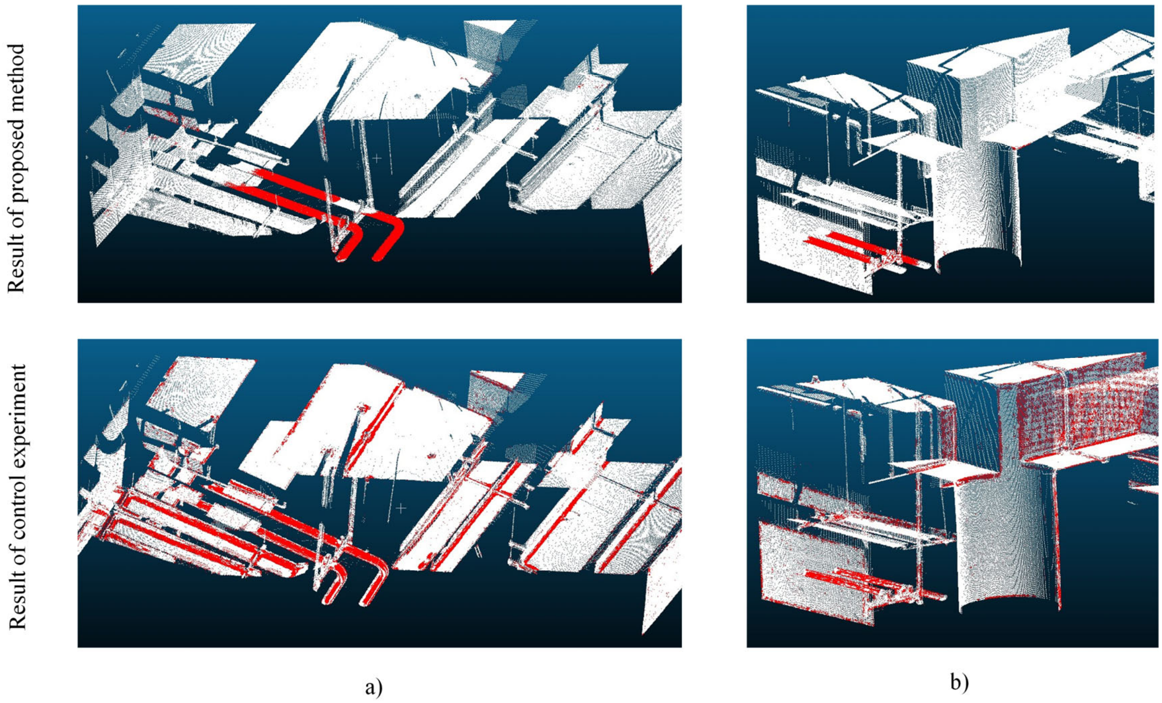

Figure 9. Visible result comparison for the intercepted datasets. (**a**) shows the result comparison on dataset 2 and (**b**) shows the result comparison on dataset 5.

According to Figure 9, the results show that the proposed method can significantly focus more on the pipeline components. This could be the result of applying UnitNet and FeatureNet to process features separately and the better combination of local features and global features carried out in FinalNet. The mean accuracy of the proposed method is over 98% and the visible results also indicate the effectiveness of the pipeline extraction process. All of these results of the experiments prove that the proposed method is workable and could be further studied and applied for MEP systems.

6. Conclusions and Discussion

In this work, we proposed a deep neural network that is designed especially for complex construction industry applications and demonstrated the effectiveness of our method. Firstly, we built a dataset based on the engineering situation of the construction industry. Then, we established a neural network structure in which local features and global features are processed separately in UnitNet and FeatureNet. We made full use of traditional shape features to enrich the features of simple point clouds and avoid excessive dependence on them through the structural design of the neural networks. Further, a feature map was proposed in FinalNet for feature fusion. Finally, by establishing a control experiment and comparing the results, the method proposed in this paper offers an effective and feasible solution for deep learning applications for classification in the construction industry.

This study makes three main contributions to the knowledge and engineering informatics. Firstly, we have optimized the combination of traditional shape features and deep learning networks, which improves the accuracy of the engineering information collected for 3D scenes and the feasibility of this method, as proven via experiments. Secondly, we have established a targeted dataset for the actual situation in construction and engineering, which provides a data basis for related research in the construction field in the future and expands the breadth of engineering information in three-dimensional space. Thirdly, we have provided efficient solutions to replace manual labor in the processes of 3D reconstruction and as-built BIM in the construction industry, since they are vital parts of the engineering information digitization process. Among these contributions, the most important is that this study provides a feasible and effective method of MEP system reconstruction and digitalization, and this method could be further studied for application in other construction fields.

Last but not least, although the proposed method achieved good performance in the experiment, there are still some problems to cover in future studies. Because the extraction of local features is based on the description of the neighboring points, the classification accuracy of the boundary part of the input data is relatively lower than other parts. Additionally, although this paper has initially established a dataset for construction engineering, there are comparatively few categories available. Regarding the experiments, the design of this work could also be improved for the simplification of input data, and the binary classification should be further extended to multiple classifications of other MEP components.

Author Contributions: Conceptualization, Z.X., R.K. and H.L.; methodology, Z.X., R.K. and H.L.; software, R.K.; validation, R.K; investigation, R.K.; writing—original draft preparation, Z.X., R.K. and H.L.; writing—review and editing, Z.X., R.K. and H.L.; visualization, Z.X. and R.K.; funding acquisition, Z.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (NSFC-72071043), Natural Science Foundation of Jiangsu Province (BK20201280), MOE (Ministry of Education in China) and Project of Humanities and Social Sciences (20YJAZH114).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are contained within the article. Additional supporting data presented in this study are available on request from the corresponding author.

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

References

1. Ma, Z.; Liu, S. A review of 3D reconstruction techniques in civil engineering and their applications. *Adv. Eng. Inform.* **2018**, *37*, 163–174. [CrossRef]
2. Wang, C.; Cho, Y.K.; Kim, C. Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Autom. Constr.* **2015**, *56*, 1–13. [CrossRef]
3. 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]
4. 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]
5. Wang, Y.; Pan, G.; Wu, Z.; Han, S. Sphere-spin-image: A viewpoint-invariant surface representation for 3D face recognition. In Proceedings of the International Conference on Computational Science, Krakow, Poland, 6–9 June 2004.
6. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic graph cnn for learning on point clouds. *ACM Trans. Graph.* **2019**, *38*, 1–12. [CrossRef]
7. Barazzetti, L. Parametric as-built model generation of complex shapes from point clouds. *Adv. Eng. Inform.* **2016**, *30*, 298–311. [CrossRef]

8. Xue, F.; Lu, W.; Chen, K.; Webster, C.J. BIM reconstruction from 3D point clouds: A semantic registration approach based on multimodal optimization and architectural design knowledge. *Adv. Eng. Inform.* **2019**, *42*, 100965. [CrossRef]
9. Gao, T.; Akinci, B.; Ergan, S.; Garrett, J. An approach to combine progressively captured point clouds for BIM update. *Adv. Eng. Inform.* **2015**, *29*, 1001–1012. [CrossRef]
10. Rodríguez-Gonzálvez, P.; Rodríguez-Martín, M.; Ramos, L.F.; González-Aguilera, D. 3D reconstruction methods and quality assessment for visual inspection of welds. *Autom. Constr.* **2017**, *79*, 49–58. [CrossRef]
11. Rodríguez-Moreno, C.; Reinoso-Gordo, J.F.; Rivas-López, E.; Gómez-Blanco, A.; Ariza-López, F.J.; Ariza-López, I. From point cloud to BIM: An integrated workflow for documentation, research and modelling of architectural heritage. *Surv. Rev.* **2018**, *50*, 212–231. [CrossRef]
12. Krijnen, T.; Beetz, J. An IFC schema extension and binary serialization format to efficiently integrate point cloud data into building models. *Adv. Eng. Inform.* **2017**, *33*, 473–490. [CrossRef]
13. Agapaki, E.; Brilakis, I. CLOI-NET: Class segmentation of industrial facilities' point cloud datasets. *Adv. Eng. Inform.* **2020**, *45*, 101121. [CrossRef]
14. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
15. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
16. Cao, M.-T.; Tran, Q.-V.; Nguyen, N.-M.; Chang, K.-T. Survey on performance of deep learning models for detecting road damages using multiple dashcam image resources. *Adv. Eng. Inform.* **2020**, *46*, 101182. [CrossRef]
17. Czerniawski, T.; Leite, F. Automated segmentation of RGB-D images into a comprehensive set of building components using deep learning. *Adv. Eng. Inform.* **2020**, *45*, 101131. [CrossRef]
18. Koo, B.; Jung, R.; Yu, Y. Automatic classification of wall and door BIM element subtypes using 3D geometric deep neural networks. *Adv. Eng. Inform.* **2021**, *47*, 101200. [CrossRef]
19. Saovana, N.; Yabuki, N.; Fukuda, T. Development of an unwanted-feature removal system for Structure from Motion of repetitive infrastructure piers using deep learning. *Adv. Eng. Inform.* **2020**, *46*, 101169. [CrossRef]
20. Hichri, N.; Stefani, C.; De Luca, L.; Veron, P.; Hamon, G. From point cloud to bim: A survey of existing approaches. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci* **2013**, *XL-5/W2*, 343–348. [CrossRef]
21. Golovinskiy, A.; Funkhouser, T. Randomized cuts for 3D mesh analysis. In Proceedings of the ACM SIGGRAPH Asia 2008, Singapore, 10–13 December 2008.
22. Kalogerakis, E.; Hertzmann, A.; Singh, K. Learning 3D mesh segmentation and labeling. In Proceedings of the ACM SIGGRAPH 2010, Los Angeles, CA, USA, 26–30 July 2010.
23. Woo, H.; Kang, E.; Wang, S.; Lee, K.H. A new segmentation method for point cloud data. *Int. J. Mach. Tools Manuf.* **2002**, *42*, 167–178. [CrossRef]
24. Guo, K.; Zou, D.; Chen, X. 3D Mesh Labeling via Deep Convolutional Neural Networks. *ACM Trans. Graph.* **2015**, *35*, 1–12. [CrossRef]
25. Guo, Y.; Wang, H.; Hu, Q.; Liu, H.; Liu, L.; Bennamoun, M. Deep learning for 3d point clouds: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 4338–4364. [CrossRef]
26. Wu, Z.; Song, S.; Khosla, A.; Yu, F.; Zhang, L.; Tang, X.; Xiao, J. 3d shapenets: A deep representation for volumetric shapes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
27. Armeni, I.; Sener, O.; Zamir, A.R.; Jiang, H.; Brilakis, I.; Fischer, M.; Savarese, S. 3d semantic parsing of large-scale indoor spaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
28. Tchapmi, L.; Choy, C.; Armeni, I.; Gwak, J.; Savarese, S. Segcloud: Semantic segmentation of 3d point clouds. In Proceedings of the 2017 International Conference on 3D Vision (3DV), Qingdao, China, 10–12 October 2017.
29. Lin, Y.; Jiang, M.; Yao, Y.; Zhang, L.; Lin, J. Use of UAV oblique imaging for the detection of individual trees in residential environments. *Urban For. Urban Green.* **2015**, *14*, 404–412. [CrossRef]
30. Liu, J.; Xu, G.; Ren, L.; Qian, Z.; Ren, L. Defect intelligent identification in resistance spot welding ultrasonic detection based on wavelet packet and neural network. *Int. J. Adv. Manuf. Technol.* **2017**, *90*, 2581–2588. [CrossRef]
31. Liu, P.; Li, Y.; Hu, W.; Ding, X.B. Segmentation and reconstruction of buildings with aerial oblique photography point clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 109. [CrossRef]
32. Ibrahim, M.; Smith, R.; Wang, C.H. Ultrasonic detection and sizing of compressed cracks in glass-and carbon-fibre reinforced plastic composites. *Ndt E Int.* **2017**, *92*, 111–121. [CrossRef]
33. Kim, M.-K.; Wang, Q.; Park, J.-W.; Cheng, J.C.; Sohn, H.; Chang, C.-C. Automated dimensional quality assurance of full-scale precast concrete elements using laser scanning and BIM. *Autom. Constr.* **2016**, *72*, 102–114. [CrossRef]
34. Shi, S.; Wang, X.; Li, H. Pointrcnn: 3d object proposal generation and detection from point cloud. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.
35. Pu, S.; Vosselman, G. Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 575–584. [CrossRef]
36. Xu, Y.; Tuttas, S.; Hoegner, L.; Stilla, U. Geometric Primitive Extraction From Point Clouds of Construction Sites Using VGS. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 424–428. [CrossRef]

37. 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]
38. Tsai, Y.-H.; Wang, J.; Chien, W.-T.; Wei, C.-Y.; Wang, X.; Hsieh, S.-H. A BIM-based approach for predicting corrosion under insulation. *Autom. Constr.* **2019**, *107*, 102923. [CrossRef]
39. Arulogun, O.; Falohun, A.; Akande, N. Radio frequency identification and internet of things: A fruitful synergy. *Br. J. Appl. Sci. Technol.* **2016**, *18*, 1–16. [CrossRef]
40. Domdouzis, K.; Kumar, B.; Anumba, C. Radio-Frequency Identification (RFID) applications: A brief introduction. *Adv. Eng. Inform.* **2007**, *21*, 350–355. [CrossRef]
41. Patil, A.K.; Holi, P.; Lee, S.K.; Chai, Y.H. An adaptive approach for the reconstruction and modeling of as-built 3D pipelines from point clouds. *Autom. Constr.* **2017**, *75*, 65–78. [CrossRef]
42. Tran, T.-T.; Cao, V.-T.; Laurendeau, D. Extraction of cylinders and estimation of their parameters from point clouds. *Comput. Graph.* **2015**, *46*, 345–357. [CrossRef]
43. Johnson, A.E.; Hebert, M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 433–449. [CrossRef]
44. Tombari, F.; Salti, S.; Di Stefano, L. Unique signatures of histograms for local surface description. In Proceedings of the European Conference on Computer Vision, Heraklion, Greece, 5–11 September 2010.
45. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [CrossRef]
46. Li, B. 3d fully convolutional network for vehicle detection in point cloud. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2017, Vancouver, BC, Canada, 24–28 September 2017.
47. Hackel, T.; Savinov, N.; Ladicky, L.; Wegner, J.D.; Schindler, K.; Pollefeys, M. Semantic3d. net: A new large-scale point cloud classification benchmark. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-1/W1*, 91–98. [CrossRef]
48. Yang, Y.; Feng, C.; Shen, Y.; Tian, D. Foldingnet: Point cloud auto-encoder via deep grid deformation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
49. Zhou, Y.; Tuzel, O. Voxelnet: End-to-end learning for point cloud based 3d object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
50. Huang, J.; You, S. Point cloud labeling using 3d convolutional neural network. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016.
51. Klokov, R.; Lempitsky, V. Escape from cells: Deep kd-networks for the recognition of 3d point cloud models. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
52. Fang, Q.; Li, H.; Luo, X.; Ding, L.; Luo, H.; Rose, T.M.; An, W. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Autom. Constr.* **2018**, *85*, 1–9. [CrossRef]
53. Fang, Q.; Li, H.; Luo, X.; Ding, L.; Rose, T.M.; An, W.; Yu, Y. A deep learning-based method for detecting non-certified work on construction sites. *Adv. Eng. Inform.* **2018**, *35*, 56–68. [CrossRef]
54. Fang, Y.; Xie, J.; Dai, G.; Wang, M.; Zhu, F.; Xu, T.; Wong, E. 3d deep shape descriptor. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
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. Sun, S.; Wang, S.; Wei, Y. A new ensemble deep learning approach for exchange rates forecasting and trading. *Adv. Eng. Inform.* **2020**, *46*, 101160. [CrossRef]
57. Ziari, H.; Sobhani, J.; Ayoubinejad, J.; Hartmann, T. Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods. *Int. J. Pavement Eng.* **2016**, *17*, 776–788. [CrossRef]
58. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Processing Syst.* **2012**, *25*, 1097–1105. [CrossRef]
59. Deng, F.; He, Y.; Zhou, S.; Yu, Y.; Cheng, H.; Wu, X. Compressive strength prediction of recycled concrete based on deep learning. *Constr. Build. Mater.* **2018**, *175*, 562–569. [CrossRef]
60. Nguyen, T.; Kashani, A.; Ngo, T.; Bordas, S. Deep neural network with high-order neuron for the prediction of foamed concrete strength. *Comput.-Aided Civ. Infrastruct. Eng.* **2019**, *34*, 316–332. [CrossRef]
61. Rahman, A.; Srikumar, V.; Smith, A.D. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl. Energy* **2018**, *212*, 372–385. [CrossRef]
62. Rafiei, M.H.; Adeli, H. Novel machine-learning model for estimating construction costs considering economic variables and indexes. *J. Constr. Eng. Manag.* **2018**, *144*, 04018106. [CrossRef]
63. Kolar, Z.; Chen, H.; Luo, X. Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images. *Autom. Constr.* **2018**, *89*, 58–70. [CrossRef]
64. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [CrossRef]
65. Pathirage, C.S.N.; Li, J.; Li, L.; Hao, H.; Liu, W.; Ni, P. Structural damage identification based on autoencoder neural networks and deep learning. *Eng. Struct.* **2018**, *172*, 13–28. [CrossRef]
66. Eastman, C.M.; Eastman, C.; Teicholz, P.; Sacks, R.; Liston, K. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*; John Wiley & Sons: Hoboken, NJ, USA, 2011.

67. Franz, S.; Irmler, R.; Rüppel, U. Real-time collaborative reconstruction of digital building models with mobile devices. *Adv. Eng. Inform.* **2018**, *38*, 569–580. [CrossRef]
68. Yi, L.; Kim, V.G.; Ceylan, D.; Shen, I.-C.; Yan, M.; Su, H.; Lu, C.; Huang, Q.; Sheffer, A.; Guibas, L. A scalable active framework for region annotation in 3d shape collections. *ACM Trans. Graph.* **2016**, *35*, 1–12. [CrossRef]
69. Shen, Z.; Ma, X.; Li, Y. A hybrid 3D descriptor with global structural frames and local signatures of histograms. *IEEE Access* **2018**, *6*, 39261–39272. [CrossRef]
70. Salti, S.; Tombari, F.; Di Stefano, L. SHOT: Unique signatures of histograms for surface and texture description. *Comput. Vis. Image Underst.* **2014**, *125*, 251–264. [CrossRef]
71. Assfalg, J.; Bertini, M.; Del Bimbo, A.; Pala, P. Content-based retrieval of 3-D objects using Spin Image Signatures. *IEEE Trans. Multimed.* **2007**, *9*, 589–599. [CrossRef]
72. Li, Z.; Zhang, L.; Tong, X.; Du, B.; Wang, Y.; Zhang, L.; Zhang, Z.; Liu, H.; Mei, J.; Xing, X.; et al. A three-step approach for TLS point cloud classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5412–5424. [CrossRef]
73. He, Y.Q.; Mei, Y.G. An efficient registration algorithm based on spin image for LiDAR 3D point cloud models. *Neurocomputing* **2015**, *151*, 354–363. [CrossRef]
74. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.