

Enhancing Point Cloud Semantic Segmentation in the Data-Scarce Domain of Industrial Plants through Synthetic Data

F. Noichl¹ | F. C. Collins¹ | A. Braun¹ | A. Borrmann¹

¹Chair of Computational Modeling and Simulation, Technical University of Munich, Arcisstr. 21, Munich, Germany

Correspondence

André Borrmann, Chair of Computational Modeling and Simulation, School of Engineering and Design, Technical University of Munich, Germany

Email: andre.borrmann@tum.de

ABSTRACT

Digitizing existing structures is essential for applying digital methods in architecture, engineering, and construction. However, the adoption of data-driven techniques for transforming point cloud data into useful digital models faces challenges, particularly in the industrial domain, where ground truth datasets for training are scarce. This paper investigates a solution leveraging synthetic data to train data-driven models effectively. In the investigated industrial domain, the complex geometry of building elements often leads to occlusions, limiting the effectiveness of conventional sampling-based synthetic data generation methods. Our approach proposes the automatic generation of realistic and semantically enriched ground truth data using surface-based sampling methods and laser scan simulation on industry-standard 3D models. In the presented experiments we use a neural network for point cloud semantic segmentation to demonstrate that compared to sampling-based alternatives, simulation-based synthetic data significantly improves mean class intersection over union performance on real point cloud data, achieving an up to 7% absolute increase.

KEYWORDS:

point cloud, semantic segmentation, laser scanning, TLS, synthetic data, training data, point cloud sampling, laser scan simulation, industrial facilities

1 INTRODUCTION

As-is model representations of built facilities are of great value for numerous applications in Architecture, Engineering, Construction, and Operations (AECO). Owners, operators, and planners rely on them for their purposes (Borrmann et al., 2018). For building operators, for example, these models enable more efficient facility management (Pärn et al., 2017). Owners and their respective planners require them to restructure, renovate and do other activities that intervene with the current situation (Hammond et al., 2014). These models might be available to some extent for newly planned and erected buildings, albeit with reduced geometric detail depending on the chosen approach and use case (Gregor et al., 2009); in many cases, no digital models are available at all (Talebi, 2014).

At a later stage in the building lifecycle, the as-is status might deviate heavily from the as-designed or as-built status due to undocumented changes because models are rarely fully updated (Volk et al., 2014). Furthermore, the older an existing facility is, the less likely it becomes that its stakeholders possess any useful digital representation – recreating a detailed, semantic 3D model entirely by hand is extremely expensive and time-consuming (Fumarola & Poelman, 2011). The research field of "Scan-to-BIM" focuses on methods that allow automating parts of this process to digitize the existing building stock (Bosché et al., 2015; Lu & Brilakis, 2019).

The as-is conditions of the built environment must be captured in the field first to provide the necessary data basis for approaches associated with Scan-to-BIM. Such data acquisition is preferably performed in 3D, using laser scanning or photogrammetry (Li et al., 2022), during the construction phase (Chern et al., 2023; Z. Wang et al., 2022) or in the context of existing projects (Tong et al., 2023; Wu et al., 2022; Zheng et al., 2022). This reality capture results in millions of points representing the object surfaces visible to the sensor in point clouds that can be very precise but are characterized by a few major shortcomings and cannot be directly used for further activities such as redesigning. Point clouds do not inherently carry semantic information, include noise, and lack closed surfaces caused by occlusions

(Walsh et al., 2013). To convert them into formats valuable to engineering, facility management, and other activities, they must be processed intelligently to enrich them with further information or create surface or volumetric models that can seamlessly be used in subsequent processing steps. Traditionally, this is a manual task conducted by trained engineers (J. F. Hullo & Thibault, 2014): Pre-processed data is filtered, cut into subsets such as slices, and finally used to create 3D models by hand that are as close to the captured point cloud as possible. This process is time-consuming, yields subjective results, and is therefore inflexible and expensive. Lu & Brilakis (2019) investigated the effort required for manual modeling of bridge infrastructure, Agapaki et al. (2018), addressing the challenge of resource intensity, identified the most critical objects and their frequency in industrial models, along with the manual effort of modeling them. In (Fumarola & Poelman, 2011), different approaches applied to several projects are presented and evaluated concerning their degree of automation and individual requirements. J.-F. Hullo et al. (2015) report on a large-scale study that reconstructed a nuclear reactor building from terrestrial laser scanning (TLS) and RGB image data, in which around 70% of the overall required time was spent on CAD reconstruction. Due to these circumstances, many attempts have been made to automate parts of the Scan-to-BIM process, with scopes ranging from volumetric models on building level (Ochmann et al., 2016) to domain-specific solutions (Smith & Sarlo, 2022). Beyond the generation of a model, typical scenarios include urban applications related to traffic (Balado et al., 2019; Esmoris et al., 2023) and structural health monitoring (Oh et al., 2017; H. S. Park et al., 2007; S. W. Park et al., 2015; Zhou et al., 2022).

The initial step of enriching the raw input point clouds with semantic information is highly labor-intensive as it requires the user to navigate large unstructured datasets to first identify and then manually separate objects and systems in 3D space. This step is also denoted as semantic segmentation and has received much attention because the underlying technical problem is relevant for those working with the built environment and in autonomous driving, geosciences, augmented and virtual reality applications, and many more. Traditional strategies to perform such

semantic enrichment rely on hand-crafted features or well-known geometric properties of object classes (L. Ma et al., 2018; Macher et al., 2017; Sharif et al., 2017); more recently, purely data-driven methods show the most promising results (Croce et al., 2021; Perez-Perez et al., 2021). However, the latter methods rely on the availability of large amounts of high-quality annotated point cloud data for method design and evaluation. The amount and quality of this annotated data are crucial to the success of automated methods of point cloud semantic enrichment. The process of manual annotation is costly and time-consuming, unambiguously pointed out by Huang et al. (2023) for the case of 2D ground penetrating radar data; for 3D point clouds, this issue is even more severe (Shi et al., 2021).

In domains related to urban scenes or indoor office environments, an increasing amount of open-source data is available for these purposes. In this research, however, we are focusing on the data-weak domain of industrial facilities. While capturing manufacturing plants and refinery scenes has become widespread for more than fifteen years (Shellshear et al., 2015), industrial owners and operators usually do not annotate large amounts of this data, let alone publish it, due to confidentiality and employee privacy issues. Consequently, there are too few annotated point clouds for training the machine-learning models for automated point cloud segmentation.

At the same time, for manufacturing plants that are subject to periodic changes due to product cycles and frequent adaptations for optimization, 3D design models have become an industry best practice for steel and plant construction (Wiendahl et al., 2015). This domain has been very active in developing and adapting digital methods with regards to 3D models for planning and operation (Gregor et al., 2009), evident, for instance, in national regulation for the standardization of 3D models used in the German car manufacturing industry (VDA, 2009). While detailed 3D models representing building structure and technical equipment can be utilized to generate synthetic ground truth data, their complex, intertwined geometry limits the value of synthetic data generated using conventional, sampling-based methods. Simulation methods for considering these specific conditions have not been investigated in this context.

This paper presents a method to generate realistic, semantically rich ground truth data based on specimens of such 3D design models by applying state-of-the-art laser scan simulation. This type of simulation is able to consider the precision and accuracy of the used sensors through equipment parameters and the complexity of the surrounding scene's layout. In doing so, the paper aims to contribute to an increase in performance of point cloud semantic segmentation for domains with no publicly available datasets by introducing this level of realism to synthetic data, thus reducing the amount of manually annotated data necessary to achieve useful results. While much less computationally expensive, conventional methods to generate synthetic data based on such 3D models fail to achieve the amount of realism necessary to depict complex scenes in the industrial domain sufficiently to learn distinctive features for semantic segmentation. We conduct an extensive experiment to validate this increased effort the presented simulation-based method brings in comparison to conventional, sampling-based generation methods; results are analyzed and discussed in detail. The results of the experiment unambiguously show the superiority of the more realistic, simulation-based method to generate synthetic training data. Thus, the presented approach is able to facilitate existing Scan-to-BIM approaches by improving segmentation results while minimizing manual work.

This study targets academics and professionals seeking practical solutions in industrial applications. It explores the potential of neural networks for point cloud semantic segmentation in domains where open datasets are unavailable. Our approach tests the hypothesis that synthetic datasets can effectively train models for industrial use. We offer comprehensive explanations for creating these synthetic datasets and conduct thorough experiments to validate their practicality in industrial contexts. Additionally, the paper extends the body of knowledge by comparing various data generation methods, highlighting their unique potentials and limitations.

The paper is structured as follows: Section 1 introduces the research topic and provides an overview of our approach. In Section 2, we review relevant prior research to establish the context and motivation for our work. Section 3 presents our methodological approach, detailing the

methods and techniques used in our study. Section 4 discusses the extensive experiments conducted and presents the obtained results. Section 5 is dedicated to the discussion of our findings and their implications. Finally, in Section 6, we conclude our presentation and provide directions for future research.

2 RELATED WORKS

This paper presents an approach to address the scarcity of training data in point cloud semantic segmentation within the industrial domain. The following subsections provide an overview of relevant related works, point cloud enrichment, domain-specific training data, and investigations into the potential of synthetically generated training data.

2.1 Point Cloud Enrichment

All options to investigate and further process the captured data of an existent structure depend on the information that can be recognized in the point cloud. If the objects of interest possess well-known geometric properties, this can be achieved using specifically chosen geometric features. One popular method is principal component analysis, which was, for example, applied in (Yokoyama et al., 2013) to detect poles in urban scenes. The well-known Stanford 3D Indoor Scene Dataset (S3DIS, introduced by Armeni et al. (2016)) was processed with a peak-gap-peak pattern to detect void spaces of walls to separate rooms. Czerniawski et al. (2016) use specific local curvature to detect pipes.

The application of data-driven methods allows algorithms to learn the critical features from annotated training data instead of exploiting prior knowledge about geometry or materials. In our approach, point cloud semantic segmentation with supervised deep learning (PCSS as per Xie et al. (2019)) was chosen as it allows us to generate point-wise class predictions. Current learning-based methods are able to capture critical features for a large number of classes at once. In further steps, this enables the targeted application of class-specific instance segmentation and reconstruction methods.

PointNet (Qi, Su, et al., 2017) and its successor PointNet++ (Qi, Yi, et al., 2017) were crucial development

steps for the discipline because the underlying architecture allowed deep learning on points directly without translating them into a structured representation like a voxel grid first (i.e., VoxNet (Maturana & Scherer, 2015)). The performance of architectures, evaluated on a few specific datasets, has constantly been improving since then. Xie et al. (2019) and Zhang et al. (2019) present recent overviews of the topic; online resources such as Papers With Code (2021) can help to provide an up-to-date roundup in this fast-changing environment. The work of Mirzaei et al. (2022) contains a comprehensive overview of methods used by state-of-the-art point cloud deep learning network architectures. Among others, a notable performance increase for semantic segmentation on point clouds was achieved by applying Kernel Point Convolutions (Thomas et al., 2019) and Point Cloud Transformers (Guo et al., 2021).

In the domain of AECO, these developments were followed with much interest, as they offer a universal first step towards a solution of the automation of the Scan-to-BIM process, compared to the application of hand-crafted features. Perez-Perez et al. (2021) introduced Scan2BIM-Net, a combination of different networks for semantic segmentation of a case-study indoor environment point cloud dataset.

Industrial facilities pose more specific challenges than commercial buildings, along with different important object classes. Yin et al. (2021) adapted the PointNet++ architecture to their version of ResPointNet++ to achieve improved results for PCSS in an industrial environment. In (Agapaki & Brilakis, 2020), PointNet++ was extended by a neighborhood consideration to enhance its performance on the authors' manually annotated industrial dataset. The same authors expanded their scope to an instance segmentation approach, starting from an ideal set of semantic segments using a search algorithm and boundary segmentation (Agapaki, 2020).

2.2 Domain-specific Training Data for Point Cloud Semantic Segmentation

As for all data-driven methods, the performance of network architectures for point cloud semantic segmentation heavily depends on the quality and quantity

of available training data (Gao et al., 2020). This issue has been addressed and partially solved for some domains with the availability of such large-scale open-source datasets. For indoor office environments, well-known examples are the aforementioned S3DIS (Armeni et al., 2016) and ScanNet (Dai et al., 2017). For outdoor urban scenes relevant to the development of autonomous driving and smart infrastructure, among others, there are the datasets of KITTI (Geiger et al., 2013), Vaihingen (Rottensteiner et al., 2013), Paris-Lille (Roynard et al., 2018), and more.

Some of the introduced works applying methods of PCSS in AECO have specifically addressed the industrial domain (Agapaki & Brilakis, 2020; Yin et al., 2021) and introduced their work along with datasets the authors prepared and used for the development and validation of their methods. To this date, there are no labeled point cloud datasets publicly available for complete industrial scenes, which has been identified as a significant bottleneck for the wider adoption of PCSS by Cazorla et al. (2021).

2.3 Synthetic Training Data for Semantic Segmentation

For applications where such ground truth data is rare, researchers and practitioners have considered circumventing the effort for manual data collection and labeling along with potential privacy issues by utilizing synthetic data. Among those approaches are attempts in 2D to generate images and depth maps from 3D models with HoliCity (Zhou et al., 2020) and to extract frames from video games (Richter et al., 2016). Similar efforts to generate synthetic images have been made by Hong et al. (2021) in the AECO domain.

To generate point clouds that exhibit realistic properties to imitate real laser scan point clouds, some approaches work on top of existing simulation tools, such as the CARLA simulator for autonomous driving (Dosovitskiy et al., 2017). This framework was used to simulate laser scan point clouds in an urban environment similar to the KITTI dataset (Geiger et al., 2013) to create the so-called KITTI-CARLA dataset (Deschaut, 2021) and similarly for the PARIS-CARLA-3D (Deschaut et al., 2021) dataset. With SynthCity, Griffiths & Boehm (2019) provide a synthetic point cloud dataset representing urban scenes along with a highly realistic, textured 3D model of the city.

To investigate the value synthetic data has for use as training data and thus the added value it can bring to Scan-to-BIM toolchains, several related contributions are relevant for this work: Frías et al. (2022) used BIM objects to generate synthetic point clouds by sampling, to then render them to images and use them for object classification. For the application in historical buildings, Morbidoni et al. (2020) used synthetic, sampling-based point cloud data generated based on structural components of available 3D models to train an adapted version of DGCNN (Y. Wang et al., 2019) for semantic segmentation. In the context of office environments, some studies (J. W. Ma et al., 2020; Zhai et al., 2022) used the S3DIS dataset to investigate the potential of synthetic point cloud data for training a neural network for point cloud semantic segmentation.

For the experiment presented by (J. W. Ma et al., 2020), a subset of the S3DIS dataset ("Area 1") was remodeled manually in an engineering application. The pure geometry of the objects in the model was then exported to sample evenly spaced points on a 3D grid within the objects' volumes to generate synthetic training data and finally annotate the point cloud on an instance level with accordingly reduced manual effort. Subsequently, this data was used to train neural networks for semantic segmentation. The study showed that an increase of 7.1% in accuracy was feasible by augmenting a small real-world dataset with an additional large set of synthetic, sampling-based data. Finally, an experiment was conducted to investigate the potential of hybrid training datasets with varying compositions between real and synthetic data in steps of 20%. In this, the dataset with 80% real and 20% synthetic data fell short only 1.52% of the best-performing real dataset in terms of accuracy. These findings, while promising, leave some questions as to the realism of the data generated using the presented method for sampling based on a volumetric grid. We presume that with more realistic synthetic data, better results can also be obtained on real data.

The above-introduced research shows the lack of suitable training data as a critical bottleneck for applying data-driven methods for point cloud semantic segmentation in specific domains beyond commonly used, unspecific benchmark environments. Beyond existing

related works, an investigation of domain-specific potential added value of synthetic data used for training networks is identified as a research gap that this paper aims to fill. The industrial domain poses specific challenges in terms of relevant object classes, and complex, intertwined object surfaces openly visible in the facilities. These boundary conditions motivate the implementation of two different methods of synthetic data generation that allow us to take into account these particular circumstances to varying extents. Complex industrial facilities often possess detailed 3D models, representing at least the as-designed status of the building structure, technical equipment in the buildings, and, depending on the domain, production equipment, logistic systems, and more. These models include complete scenes and can be leveraged to create synthetic point cloud data that can be used to train networks for point cloud semantic segmentation. Depending on the chosen method and on the quality and structure of the available models, the data generation process can be fully automated or achieved with very little manual effort: Large amounts of annotated point cloud data can be generated automatically to train and improve data-driven methods for flexible requirements for respective 3D models of industrial facilities.

3 Research Methodology

This paper investigates to which extent synthetic point clouds of varying quality can be used to address the challenge that the insufficient availability of annotated training data poses to the applicability of data-driven approaches for point cloud semantic segmentation for building infrastructure in the industrial domain.

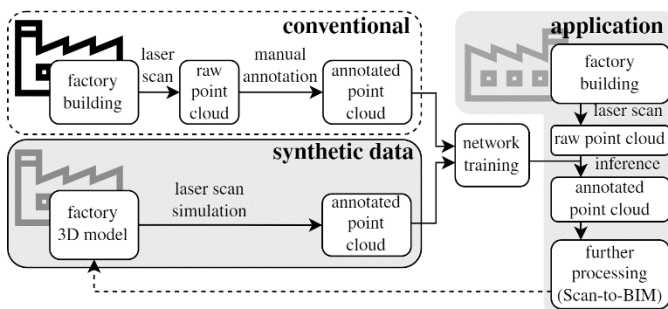


FIGURE 1. Data preparation workflow, highlighting differences between conventional, manual annotation and the generation of synthetic, annotated data using existing 3D models.

The manual effort otherwise required to annotate such datasets can be completely avoided or reduced drastically while avoiding human errors in the annotation process. The only fixed requirement persists in the input 3D models, which are, in many cases, readily available. **FIGURE 1** depicts the underlying logical structure of the presented method. The quality and inherent value of the resulting, synthetically generated point clouds are highly dependent on the quality of the underlying 3D models, especially regarding the level of detail and completeness- and the chosen method for the point cloud generation.

This paper aims to investigate the value of different types of synthetic data as training data for point cloud semantic segmentation applications in direct comparison. Two ways to generate synthetic, annotated point cloud data based on 3D models of industrial facilities are introduced and applied to an illustrative model; two independent reference datasets are collected by terrestrial laser scanning in industrial facilities and manually annotated to enable the evaluation of real-world applicability and generalization potential. Subsequently, several semantic segmentation experiments are performed using a fixed training and testing setup for all mentioned datasets as training and test data, respectively. The final evaluation is performed with regard to the real laser scan datasets to show the extent of actual industrial applicability.

Beyond the applicability of homogeneous, synthetic training data, in a second experiment, the purely synthetic datasets are combined with small subsets of real data to investigate the potential value that can be achieved with minor manual annotation. This combination of synthetic data with real-world data is expected to improve results and is therefore referred to as *augmentation* in the following. While this implies an increased effort for data preparation, it is a handy solution to achieve results with no significant shortcomings.

3.1 Point cloud datasets

The objective of the conducted experiments is to assess the practical applicability of synthetic data within the context of industrial facilities. To achieve this, we employ suitable reference datasets in the form of real laser scan

point clouds captured using industry-standard TLS equipment.

There are multiple solutions to create a synthetic point cloud based on a given 3D model. In this paper, two alternatives are presented and compared: one based on sampling and one using state-of-the-art laser scan simulation. Both methods allow for the preservation of the semantics of the model in the process such that the resulting point cloud is fully annotated and can directly be used to train a neural network model for semantic segmentation. The described steps are chosen such that the procedure can be applied based on any conventional 3D model. No color or material attributes are used in either approach. While they would help to improve model performance (Zhai et al., 2022), standard industrial 3D models do not commonly contain this information: to ensure industrial applicability with no overhead effort for data preparation, only the model geometry is taken into account for data generation. **FIGURE 2** depicts an overview of the process of generating the synthetic point clouds as described in the following.

In the first step, the modeled scene must be split into separate entities representing semantic classes or instances. All objects are first exported separately into individual OBJ (Wavefront OBJ) files to comply with a pre-defined class split according to the classes that should be included in the semantic segmentation. The effort necessary for this step depends on the model structure in terms of semantics and how well they can be mapped to the classes that should be investigated in the point cloud. In a single-layered 3D CAD model, this step has to be performed manually: connected objects that include multiple classes might have to be separated. If the model contains all necessary information for this split and is, for example, stored in the Industry Foundation Classes (IFC) format (ISO, 2018), this can be fully automated by parsing relevant object properties. The further steps undertaken differ between simulation and sampling.

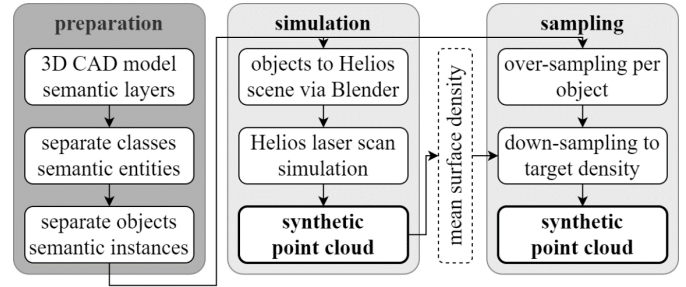


FIGURE 2. Steps for the process of generating the simulation- and sampling-based synthetic point clouds: Model data preparation (left), simulation-based data generation (middle) and sampling-based data generation (right).

Multiple solutions are available to simulate a laser scan on a given 3D geometry. As the resulting data should resemble an actual laser scan, equipment, and measurement behavior should be emulated as well as possible; model semantics must be included in the resulting point cloud to avoid any manual annotation effort. The measurement itself is based on a line-of-sight evaluation between a sensor emitting laser rays in patterns according to equipment-specific functionality and the surface of an object in the scene. Depending on distance, incidence angle, surface material parameters, and equipment parameters such as precision, simulation engines return s close to real laser scans.

Existing solutions for this include the educational platform of VRScan3D (Luhmann et al., 2022), the Blender-based (Blender Online Community, 2021) tool of BlenSor with a focus on mobile scanning platforms and depth-cameras (RGB-D) (Gschwandtner et al., 2011) and Helios++ (Winiwarter et al., 2021), with a wide variety of applications and flexible setup opportunities including different sensor types and mobile platforms. For this context, Gonzalez Stefanelli et al. (2022) present an overview of suitable platforms for data generation based on 3D building models. In our approach, we use Helios++ for laser scan simulation. The simulation kernel is based on ray tracing, simulating laser beams by sampling from probabilistic distributions, and considering material-specific reflectivity parameters.

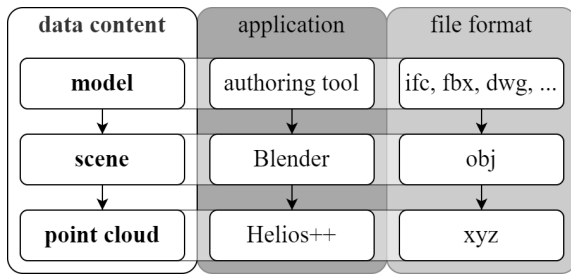


FIGURE 3. "BIM-to-Scan" workflow, in terms of data content, application, and file formats used, adapted from Noichl et al. (2021).

The process of generating synthetic, annotated point clouds based on semantic 3D models through laser scan simulation is depicted in **FIGURE 3**. As a primary step for data preparation for laser scan simulation, the 3D model is exported from the authoring tool using conventional CAD exchange formats (*.fbx, *.dwg) or *.ifc format. All of these exchange formats can be imported into Blender by default or using specific add-ins¹. After import into Blender, OBJ objects are collected in a scene in Blender to a Helios simulation scene by using our adaption of the Blender2Helios tool (Neumann, 2020). The simulation tool itself allows for customizing all aspects of the simulation. In a set of XML files, the user can define scanner properties such as range, resolution, precision, and field of view.

Furthermore, scanning locations in the survey are specified, as well as the scene itself, which is built from a set of geometric objects stored in separate OBJ files. The simulation itself takes into account these equipment parameters to simulate rays cast from the virtual scanner's sensor location, trace them, and report intersections with the scene as hits. Therefore, any point in the resulting point cloud can be clearly attributed to the class information from the underlying object. This depicts a perfect, error-free annotation as part of the process that is reproducible and scalable (Winiwarter et al., 2021). The user-definable parameters include the field of view, resolution, and precise coordinates of the laser scan sensor in the scene. The simulated point cloud possesses realistic properties, such as occlusions and minor measurement inaccuracies. After the simulation step is complete, we calculate the mean surface density of the simulation-based synthetic point cloud for later use in creating the sampling-based synthetic data (cf. **FIGURE 2**).

While laser scan simulation produces realistic results, it is also computationally expensive. Point clouds can be generated directly on the previously prepared parts of the 3D model as we process them as a triangulated mesh in OBJ format. There are various methods to do so: As they describe the surface, triangle vertices can be directly interpreted as points of the point cloud. Depending on mesh resolution and face size, this can lead to sparse clouds and highly irregular point densities. To achieve a more uniform point distribution, points can be randomly sampled on each face's surface, with the number of points per face determined by the face's area. Poisson disk sampling (Corsini et al., 2012) is an alternative method that is able to distribute points on the faces of the triangulated mesh representation of the model even more homogeneously. These methods are common practice in the field and are implemented in widely used open-source tools like CloudCompare (CloudCompare, 2021) or MeshLab (Cignoni et al., 2008).

We start the sampling process by first over-sampling points for each semantic object. Based on the initial class split, a manually defined, high number of points is sampled on the surface of the class objects using the random sampling functionality of CloudCompare (CloudCompare, 2021). The number of points is chosen so that the resulting surface density exceeds the mean surface density from the simulation. By doing so, the full surfaces of all meshes are covered with points, regardless of location or orientation. Surfaces within a model that are either contained within other objects or located within other geometric bodies, like layers within walls or ceilings, are incorporated into the resulting point cloud.

After sampling, the surface density is calculated based on the surface geometry to ensure the chosen amount of points is sufficient to obtain the mean surface density of the simulation-based point cloud. Otherwise, sampling is repeated with an increased number of points until the target is met. The over-sampling step is necessary because it is impossible to sample points to generate a specific point density without prior calculations on the underlying geometry. After a sufficient point density has been verified,

¹ such as <https://blenderbim.org/>

the over-sampled point cloud is down-sampled to ensure the same minimum point spacing as in the simulation-based point cloud (cf. **FIGURE 2**).

The resulting point clouds have the same mean surface density but different overall properties. **FIGURE 4** illustrates these differences in properties concerning local point densities and occlusions in a simple example.

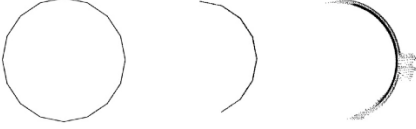


FIGURE 4. Exemplary pipe cross-section segment from three different point cloud sources: Synthetic sampling-based (left), synthetic simulation-based (middle), and real laser scan (right).

3.2 Semantic Segmentation

The core steps of our experiments are the training, testing, and evaluation of a state-of-the-art neural network for point cloud semantic segmentation on variations of our data. For semantic segmentation, the method of Kernel Point Convolutions (KPConv) (Thomas et al., 2019) is currently among the best-performing convolution methods.

For network architecture, we therefore use the Kernel Point Fully Convolutional Neural Network (KP-FCNN), which is a fully convolutional network for semantic segmentation introduced by the authors along with KPConv (Thomas et al., 2019). This architecture is well established, has been used in related studies (Deschaud et al., 2021; Soilán et al., 2021), and is among the best-performing architectures for semantic segmentation on S3DIS (Papers With Code, 2021), a core benchmark for point cloud semantic segmentation for indoor scenes in the built environment. The presented work uses the available Pytorch implementation, as published on Hugues Thomas' public GitHub repository (Thomas, 2021).

This work focuses on investigating the value of different types of point cloud data used for training a neural network for semantic segmentation. The specific performance in question is the trained network's ability to correctly predict class labels per point in a real laser scanning point cloud test set that is not used in the training phase. To do so, one and the same neural network architecture is trained on a variety of datasets of synthetic, real, and hybrid point clouds from scratch. After running inference on the real laser scanning point cloud test set, these results are compared to the

ground truth class labels of the test set. In this final evaluation step, the metrics of F_1 -score and Intersection over Union (IoU) are investigated. These metrics can be calculated based on correct predictions (denoted as true positives TP and true negatives TN) and false predictions (denoted as false positives FP and false negatives FN) as follows:

- Precision:

$$P = \frac{TP}{TP + FP} \quad (1)$$

- Recall:

$$R = \frac{TP}{TP + FN} \quad (2)$$

- F_1 -score:

$$F_1 = 2 \times \frac{P \times R}{P + R} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (3)$$

- Intersection over Union IoU :

$$IoU = \frac{TP}{TP + FP + FN} \quad (4)$$

For evaluating results in point cloud semantic segmentation, Intersection over Union or Jaccard Index (Equation 4) is commonly used as a measure of similarity between the ground truth point cloud dataset and the predicted point classes. While IoU is a measure of similarity between the ground truth and the prediction labels, the F_1 -score depicts the harmonic mean of precision and recall as a balanced measure of accuracy. Both metrics are evaluated at the class level to prevent skewed results that overestimate performance in imbalanced datasets. To evaluate overall experiment performance, this is implemented by first calculating the metric per class and subsequently the mean over all classes of the dataset per Equation 5, with C as the overall number of classes and μ as a placeholder of the respective investigated metric (cf. equations 3, 4), $m\mu$ indicates the calculation of a mean class metric.

$$m\mu = \frac{1}{C} \sum_{class=0}^C \mu_{class} \quad (5)$$

Furthermore, a variety of confusion matrices is evaluated to identify specific patterns of misprediction between certain classes. Instead of absolute values, confusion matrices are presented in normalized form in order to keep the content clear, facing imbalanced data.

4 EXPERIMENTS AND RESULTS

Several experiments were performed in the framework of this contribution using a state-of-the-art network for point cloud semantic segmentation on varying training- and test datasets. The difference between each experiment is mainly in the training dataset used. The individual datasets are presented in more detail below.

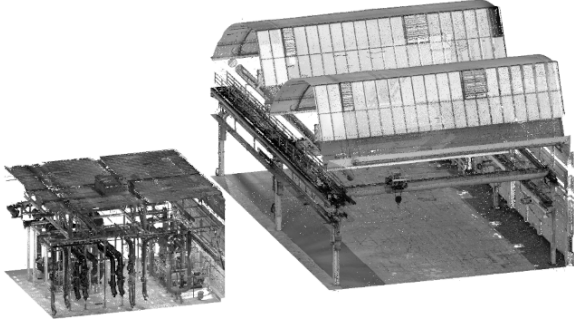


FIGURE 5. Investigated laser scan datasets, exemplary sections of the cooling plant (left) and factory hall (right), equal scale.

4.1 Datasets

Two real industrial facilities are part of these experiments: an active industrial cooling plant and a cleared factory hall. These facilities were chosen because they contain all typical objects for industrial buildings, including steel beams, pipe runs, ventilation ducts, and cable routing. Nevertheless, they differ significantly from each other. The cooling plant has a footprint of roughly 640m^2 , comprising a built volume of 4480m^3 ; the factory hall is significantly larger than the cooling plant, with a footprint of 2850m^2 and a built volume of roughly 36200m^3 . The cooling plant facility represents the core of this investigation, as it allowed performing the case study to its full extent: A detailed as-designed 3D model of the facility was available, along with the actual access to the facility to conduct a laser scan. Thus, it was possible to create point clouds reflecting this facility in the three independent ways introduced in Section 3.1: two synthetic point clouds using sampling and simulation and an actual laser scan depicting the real as-is situation in the facility.

To investigate how the findings of the core case study hold in transfer to a different exemplary dataset in a typical industrial use case, the same laser scanning system as in the cooling plant to collect point cloud data from a cleared factory hall before it was repurposed.

TABLE 1. Point cloud classes with corresponding integer ID.

ID	class name
0	wall
1	floor
2	ceiling
3	beam
4	railing
5	pipe fitting
6	pipe accessories
7	ventilation duct
8	cable routing
9	bracing
10	tank
11	clutter
12	equipment
13	noise

The classes used to annotate those point clouds are introduced in **TABLE 1**. These classes do not follow a use-case-specific structure but represent the major object types present in the case study facilities.

To gather the real laser scan datasets, a TLS scan was performed with the help of a professional surveying expert inside the case study cooling plant and the cleared factory hall. The cooling plant is the core case study dataset; the factory hall is a highly different facility yet comparable in terms of present classes. Excerpts of the two point clouds are depicted in **FIGURE 5** to illustrate this difference.

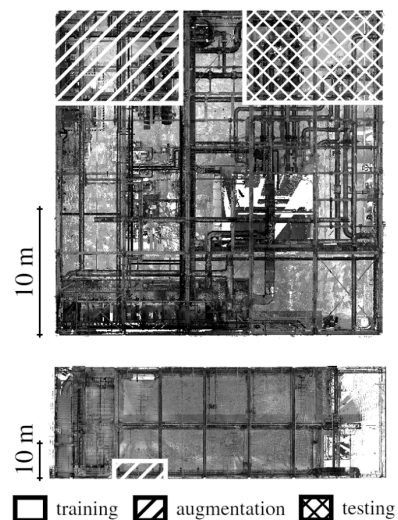


FIGURE 6. Investigated datasets, dataset splits indicated: cooling plant (top) and factory hall (bottom), 10m for scale.

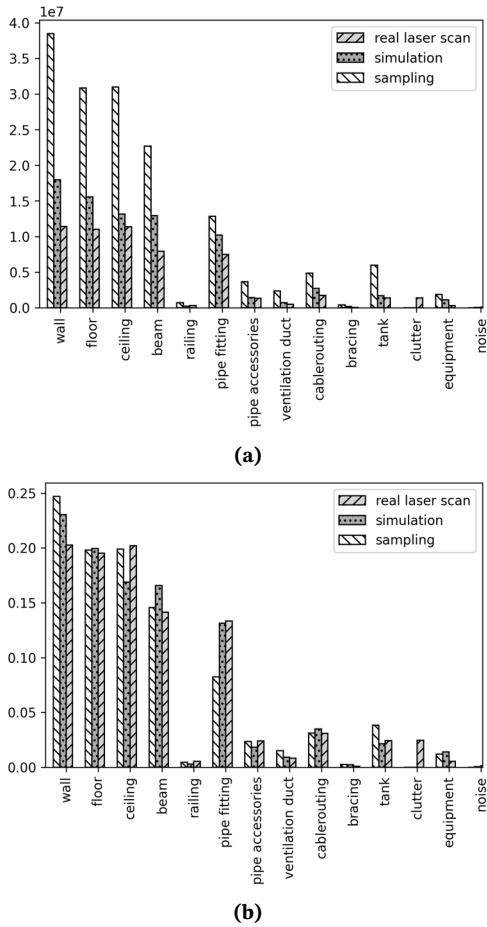


FIGURE 7. Distributions of points per class in comparison between dataset types in cooling plant: (a) total points per class and dataset type (b) normalized share of class per dataset type.

In the cooling plant, a total of 28 single scans were conducted using a FARO Focus S laser scanner, registered and processed through the manufacturer's native processing software; targetless registration could be performed as the surveyor had ensured sufficient overlaps. The resulting, registered, de-noised point cloud comprises 7×10^8 points (cf. **FIGURE 6**). For the factory hall, a total number of 15 single scans was conducted at a higher resolution than in the cooling plant to ensure sufficient density for the larger required scanning distances, which led to a total number of 6.7×10^8 points. The cooling plant point cloud was down-sampled with a minimum distance of 5mm and the factory hall to a distance of 10mm between points to create a more uniform point density throughout the datasets and reduce overall size for further processing. In the factory hall, the minimum point spacing had to be increased to keep the final resulting point cloud in a manageable size. Consecutively, the data was divided into even, box-shaped segments and down-sampled to 5×10^5 points per segment to be fit for processing a web-based

annotation tool. Manual labeling was then conducted with the AWS SageMaker GroundTruth tool (Amazon Web Services, 2021) with the classes introduced in **TABLE 1**. Annotation was performed by hand and took a total of 82 hours to complete for the cooling plant dataset, and 57 hours for the factory hall due to the simpler overall structure and experience gained from the first dataset.

After this, the manually collected label information was realigned with the input point clouds. Subsequently, the annotated points' class information was extrapolated to the points not present in the working sets after down-sampling using a k-nearest-neighbor-approach ($k = 5$). As a result, the real laser scan datasets are fully annotated with their original number of points.

Generating the synthetic datasets started from one industrial facility 3D model for both versions of data generation. The 3D CAD model used for this experiment depicts the as-designed status of the cooling plant facility. Semantic information on the contained objects is therefore organized in layers according to the responsible crafts involved in the construction project. Starting from this 3D CAD model, all objects were exported separately into individual files according to their layer-based semantics. Subsequently, the resulting collections of objects were further split or combined to comply with the pre-defined class split as introduced in **TABLE 1**.

The resulting OBJ files were then processed in different ways for the methods of simulation and sampling. For the simulation version, the sensor parameters and scanning locations in Helios were defined to be identical to those of the real TLS scan to keep the result as close to reality as possible. These results, a comparison to the real laser scan and the exact scanning parameters used are presented in Noichl et al. (2021); the simulation took around one hour to complete. In the final step, the simulation-based point cloud was combined from the single scans and subsequently down-sampled to a minimum point spacing of 5mm like the real laser scanning point cloud. The mean surface density in the resulting point cloud was calculated as roughly 25000 pts. in a radius of 5 cm using CloudCompare. For the sampling method, the separate OBJ files were processed as described with CloudCompare (cf. **FIGURE 2**). Over-sampling was performed with 10^6 points

per object to achieve sufficient density. Subsequently the resulting individual point clouds were down-sampled to the required mean surface density of 25000 pts in a 5cm radius. For larger objects where down-sampling would not reduce the absolute number of points, we repeated the process with increasing initial point numbers until this requirement was fulfilled. This iterative process took less than 10 minutes in total to compute. The separate point clouds were then combined and down-sampled with a minimum point spacing of 5cm, like both other datasets. Note that neither generation method requires specific registration as the global object coordinates are known from the beginning and preserved through the process- hence, these individual point clouds can be combined without further computation.

After the preparation of the full point clouds, all three cooling plant datasets are split into training, augmentation, and test sets. The distinct augmentation set is used for training the network and depicts the sets of data to replace and be replaced with data from the real laser scanning datasets to investigate the value of hybrid, augmented point cloud data. To create distinct datasets, the inliers of two bounding boxes in two corners of the cooling plant point clouds are separated from the full point clouds. The remaining major parts of the cooling plant point clouds are used as training sets. As the primary purpose of the factory hall dataset is testing the ability to generalize, only a small part of the dataset is separated for data augmentation purposes; the rest remains as the testing dataset. As an overview of the point cloud data used in the experiment, the real laser scan dataset of the cooling plant is depicted in **FIGURE 6**, along with an illustration of the introduced dataset splits for the cooling plant and factory hall datasets. The resulting total and per set point numbers are summarized in **TABLE 2**, which shows the variation between the data types and datasets.

TABLE 2. Dataset splits per data source type.

dataset	cooling plant						factory hall	
	sampling		simulation		laser scan		laser scan	
	points size	share	points size	share	points size	share	points size	share
	[10 ⁶]	[%]	[10 ⁶]	[%]	[10 ⁶]	[%]	[10 ⁶]	[%]
train	127.6	74.7	56.4	72.5	41.0	72.8	-	-
aug.	20.9	12.2	10.1	13.0	7.2	12.8	1.2	3.0
test	22.4	13.1	11.3	14.5	8.1	14.4	38.2	97.0
total	170.9	100.0	77.8	100.0	56.3	100.0	39.4	100.0

The actual objects vary between the as-designed and as-is state of the cooling plant facility; therefore, the class split varies between the individual datasets. Although to a lesser degree, the numbers also differ between the two synthetic datasets due to occlusions – further amplified by modeling details in the original CAD models. The collected numbers of points per class and dataset are collected in **FIGURES 7a** and **7b**. As depicted in **FIGURE 7a**, the distribution of points over the various classes is quite imbalanced, which is challenging for applying machine learning algorithms but very common for point cloud scenes, as can be seen in similar experiments (J. W. Ma et al., 2020; Soilán et al., 2021). The class analysis is omitted for the factory hall dataset to keep the study concise. While other datasets and types that aim for classification tasks (2D and 3D) can be extended by more samples of specific classes to reduce class imbalance, datasets depicting full scenes for semantic segmentation (2D and 3D) cannot be balanced easily. Publicly available datasets such as S3DIS (Armeni et al., 2016) and KITTI (Geiger et al., 2013) show similar characteristics.

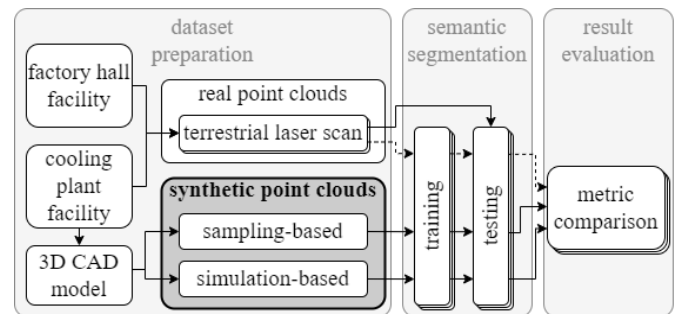


FIGURE 8. Experiment workflow for experiments 1 and 2 with regard to data sources and process steps.

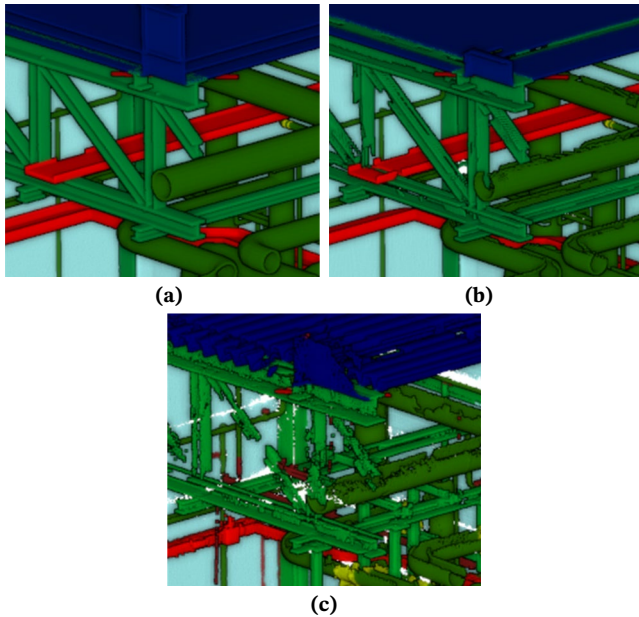


FIGURE 9. Comparison between point cloud snippets: (a) sampling-based, (b) simulation-based and (c) real, laser scanned point cloud, semantic classes of ground truth color-coded.

However, the normalized evaluation in **FIGURE 7b** shows that the overall point distribution per class is comparable throughout the datatypes. The overall dataset split is, therefore, comparable.

The sampling-based dataset has the overall highest number of points. In this dataset, all model surfaces are covered in points; neither occlusions nor model parts within convex volumes are spared. However, the difference between simulation-based and real laser scan data is not that significant. In the simulation, occlusions are realistically considered, as investigated in (Noichl et al., 2021), and depicted in **FIGURE 9**. With this, the explicitly stated limitation of sampling-based approaches, as identified in the study presented in (J. W. Ma et al., 2020), which used a volume-based sampling approach, is fulfilled. Still, the overall number of points in the simulated point cloud is slightly increased compared to the actual laser scan point cloud, as the used model is incomplete regarding highly complex surfaces and temporary and movable objects in the existing facility. First, the class of "noise" is only present in the real dataset. The synthetic datasets are generated using an as-designed model of the facility and therefore inherently do not contain non-essential or temporary objects. As the applied pre-processing includes a minimum distance down-sampling step, the reduced surface complexity leads to a reduced number of remaining

points. Furthermore, after capture, the real laser scan point cloud was cleaned of noise resulting from the facility's highly reflective materials, below-minimum distance surfaces, and incidence angles.

4.2 Experiment Setup

Two separate experiments were conducted to evaluate the validity of the proposed approach of using synthetic point cloud data for training the KP-FCNN architecture for point cloud semantic segmentation. The main steps of this experiment were repeated for

- three homogeneous datasets; two synthetic point clouds, inclusive of one sampling-based and one simulation-based version, and one real laser scanning point cloud (**Experiment 1**)
- four hybrid datasets of synthetic data augmented with a fixed amount of real laser scanning data, two from the cooling plant and the factory hall datasets each (**Experiment 2**)

To describe the steps of experiment 1, **FIGURE 8** provides an overview of the process: Data preparation for all three datasets, semantic segmentation, and evaluation of results. The machine learning model parameters remained unchanged through all trials to avoid any distortion between single experiments. For the processing in the KP-FCNN, point cloud data was pre-processed in the first step by down-sampling using a voxel grid with a consistent voxel size of 0.02m. The radius of the Kernel for convolutions to be applied on the points of the point clouds was set to 1.5m. The learning rate was fixed to 0.01, the batch size to 6, and the maximum number of epochs to 500. After training was complete, the networks were tested on the designated testing parts of the real laser scan point clouds. Thus, it was possible to investigate how well the network performed on this real test data after training on each specific training set. Subsequently, the evaluation metrics introduced in section 3.2 were calculated for each run and finally compared between experiment runs.

4.3 Experiment 1: Homogeneous training data

The first experiment investigates how well synthetic point cloud data can be used as training data for a neural network to perform point cloud semantic segmentation on

a real laser scanning point cloud. More precisely, homogeneous sets of sampling-based and simulation-based synthetic data were used to enable a direct comparison and quantify their value for application on real-world problems. The introduced network architecture was trained and tested separately on all datasets introduced above. As depicted in **FIGURE 10**, for both synthetic data runs, the loss stabilized around 350 epochs; the chosen 500 epochs of learning seem suitable for this task.

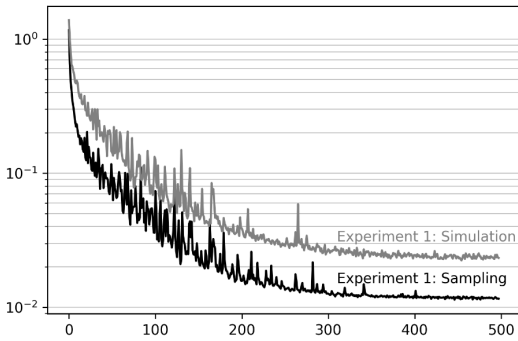


FIGURE 10. Loss curves for 500 epochs training purely on synthetic data; sampling-based and simulation-based.

As performance on synthetic data was not the purpose of this investigation, the performance of these trained networks on the designated test datasets was tested in the next step. As introduced in Section 4.1, this data was extracted from the real laser scan point clouds and was not used in training any of the networks. The test set point clouds with predicted class labels were then evaluated against the manually created ground truth. The resulting mean metrics were calculated over all classes per Equations 1 - 5 and are depicted in Table **TABLE 4**.

TABLE 3. Results for Experiment 1: Mean class Intersection over Union (mIoU) and mean class F_1 -score (mcF1) for varying setups, training on cooling plant datasets: synthetic data generated by means of sampling (SAM) and simulation (SIM); real laser scanning data (REAL) for reference.

test dataset	cooling plant			factory hall		
	SAM	SIM	REAL	SAM	SIM	REAL
mIoU	0.23	0.30	0.69	0.15	0.22	0.26
mcF1	0.30	0.38	0.75	0.20	0.28	0.36

For the cooling plant test set, the network trained on sampling-based data achieved a mean class Intersection over Union (mIoU) of 23%, while the one trained on

simulation-based data reached 30%. The network trained on the real laser scan data significantly outperformed both with a mIoU of 69%. In terms of mIoU, the simulation-based dataset outperformed the sampling-based by seven percentage points, amounting to an increase of 30% relative to the sampling-based performance. For the mean class F_1 -score (mcF1), the relative improvement averaged 27%.

For the factory hall dataset, the results ranged lower but showed the same characteristics. The model trained on real data yielded the highest results (mIoU 26%, mcF1 36%), noticeably less than the cooling plant. The type and dimension of objects in the cooling plant and factory hall datasets vary significantly. Despite this, the similarity between the separate training and testing sets within the same dataset is high. For sampling- and simulation-based training data, the difference between the datasets is less evident, resulting in lower but comparable results for the factory hall. For those, the difference in value between sampling- and simulation-based data was identical to the one measured for the cooling plant dataset, with an absolute increase of 7% mIoU and 8% mcF1.

TABLE 4. Results for Experiment 1 on cooling plant data: Class-wise F_1 -score for training on homogeneous datasets generated by means of sampling (SAM) and simulation (SIM); real laser scanning data (REAL) for reference.

	SAM	SIM	REAL
wall	0.72	0.93	0.97
floor	0.97	0.93	0.99
ceiling	0.02	0.09	1.00
beam	0.67	0.77	0.96
pipe fitting	0.53	0.57	0.92
pipe accessories	0.46	0.55	0.89
ventilation duct	0.00	0.00	0.63
cable routing	0.48	0.49	0.95
bracing	0.00	0.07	0.93
tank	0.04	0.28	0.32
clutter	0.00	0.00	0.82
equipment	0.36	0.42	0.91
noise	0.00	0.00	0.24

For the cooling plant test data, **TABLE 4** presents the F_1 -score for each class. Among the two synthetic candidate datasets, the simulation-based dataset outperforms the sampling-based dataset in most classes, with the differences

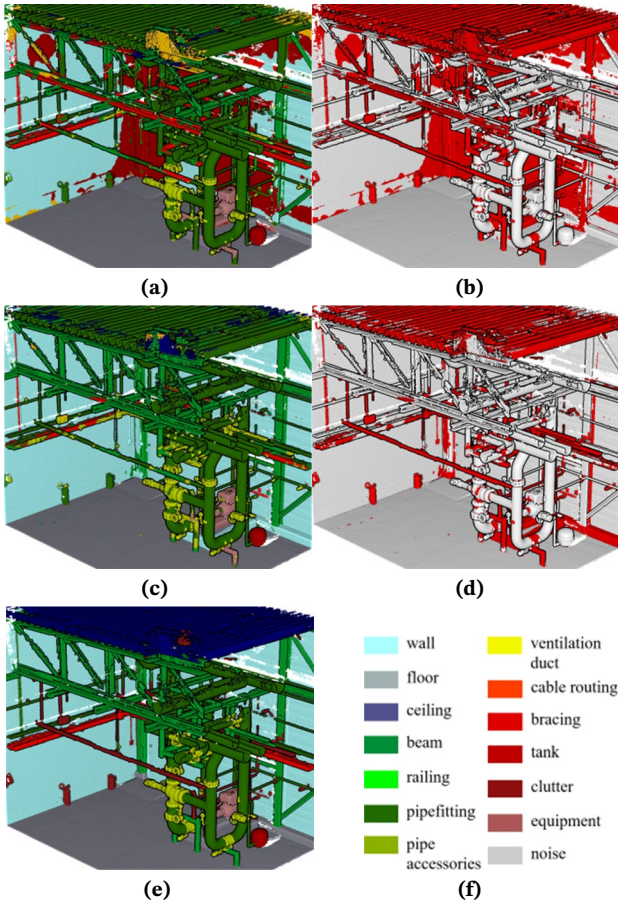


FIGURE 11. Experiment 1: Point class predictions, network trained on sampling-based (a) and simulation-based (c) datasets; False predictions (red) for training on sampling-based (b) and simulation-based (d) datasets; Ground Truth class labels (e) and legend for a, c, e (f).

ranging between 1% and 10%. However, there are a few exceptions. The ventilation duct and clutter classes both have 0% F₁-score because they are not properly represented in the synthetic dataset, and clutter is absent from all synthetic datasets. Moreover, the wall and tank classes show significant improvements of +21% and +24%, respectively, due to the discrepancy between their surface representation in the 3D CAD model and the visible surfaces in the facility. For example, the walls are modeled with several layers that are all used for sampling but are inherently occluded in the laser scan simulation. Conversely, the sampling-based dataset performs 4% better for the floor class. Although this study's findings cannot fully explain this specific exception, the overall results strongly favor the simulation-based dataset over the sampling-based dataset.

FIGURE 11 shows a snippet of point predictions per experiment run along with visualized false predictions.

Well-performing and failing classes can be distinguished as follows: *wall*, *floor*, *pipe fittings*, and *pipe accessories* range above 50%. The visual check and false prediction figures of **FIGURE 11** indicate this. The confusion matrices for Experiment 1 are depicted in **FIGURE 12**. A clear result in comparing both experiments is the inability to produce good results for elements that are not or are poorly depicted in the utilized 3D models. In both trials of Experiment 1, the facility's ceiling is poorly predicted (cf. **FIGURES 11** and **12**).

As shown in **FIGURES 8a** and **8b**, the ceiling is depicted by a simple plane in the underlying as-designed model.

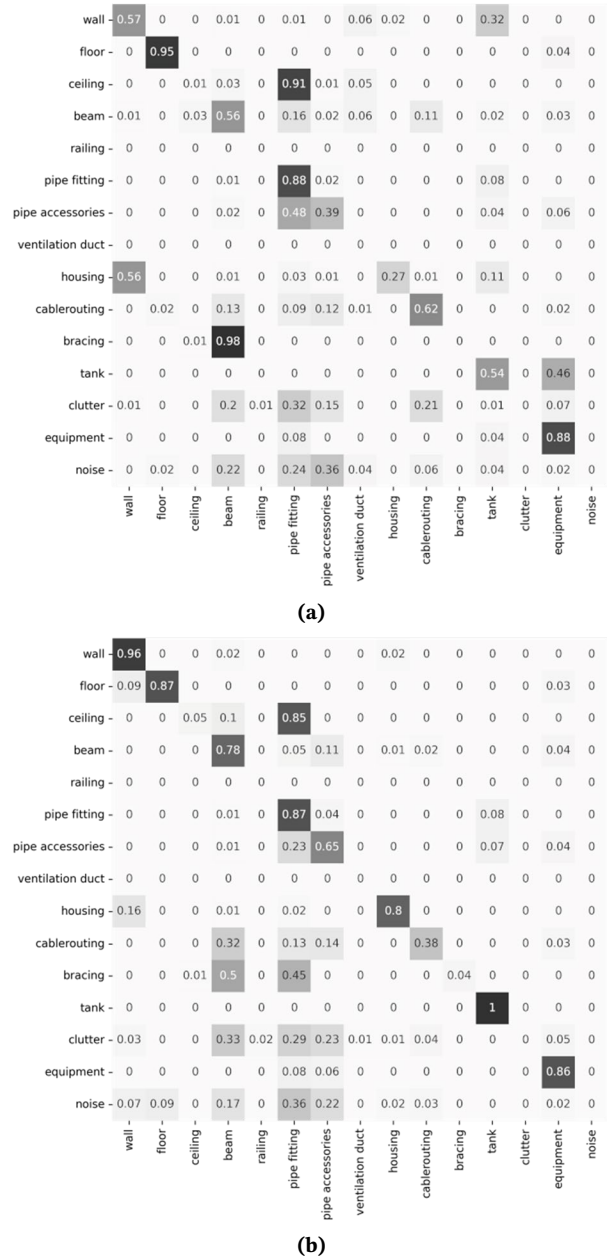


FIGURE 12. Confusion matrices of the real laser scanning test set for models trained on (a) sampling-based and (b) simulation-based synthetic training data.

While this representation might be sufficient for planning purposes, it lacks the geometric precision needed to produce synthetic point clouds for training a neural network for semantic segmentation.

The bracing elements themselves are modeled in detail, but the proximity to the ill-represented ceiling class leads to poor results for this class as well. For the sampling case, predictions for pipe accessories are mixed between pipe fitting and pipe accessories, while in the simulation-based trial, the predictions are more homogeneously pipe accessories.

4.4 Experiment 2: Augmented training data

To improve the network’s performance, the purely synthetic training point cloud data was augmented with smaller amounts of annotated, real laser scan data. In a practical application, these conditions could be achieved by generating a large amount of synthetic data leveraging industrial facility 3D models; in addition to that, a small amount of point cloud data from the facility in question or a similar one would be manually annotated and added to the synthetic data to form a hybrid training dataset.

In this experiment, the designated augmentation part of the synthetic datasets was replaced by the respective augmentation counterpart of the real laser scan point clouds for this second experiment. Thus, the underlying dataset split was not changed (cf. **FIGURE 6**) while aiming to overcome the shortcomings identified in the first experiment. For the reasons laid out in section 4.1, the total number of points varied between the datasets. Hence, the augmentation set of 7.2×10^6 points or 12.8% of the real laser scan dataset led to a different percentage of the overall available points for training in the respective augmented dataset set. For the sampling-based dataset, the substituted augmentation set constitutes 5.3% of the available points for training in the simulation-based version 11.3%. The underlying calculations were based on the dataset-specific numbers of points and relative splits introduced in **TABLE 2**. Except for the datasets used, all parameters remained unchanged from Experiment 1. Similar to the first experiment, training loss stabilized around 350 epochs; therefore, the copied training parameters were acceptable for the second experiment.

TABLE 5. Results for Experiment 2 cooling plant dataset: Macro mean class Intersection over Union (mIoU) and mean class F_1 -score (mcF1) for varying setups, training on synthetic datasets generated by means of sampling (SAM+) and simulation (SIM+) (88%) augmented by 12% real laser scanning data; 100% real laser scanning data (REAL) for reference; Δ_{SAM} , Δ_{SIM} indicate the absolute changes in comparison to Experiment 1.

	SAM+	Δ_{SAM}	SIM+	Δ_{SIM}	REAL
mIoU	0.60	+0.37	0.65	+0.35	0.69
mcF1	0.65	+0.35	0.72	+0.34	0.75

Both trained networks were again tested on the test set taken from the real laser scan dataset. Same as in Experiment 1, this data was not used in any training. As a result, the performance increases significantly for both configurations. In this setup, the augmented sampling-based data achieves 60% mIoU, and the simulation-based data even goes as high as 65%, which is close to the performance of the complete real training dataset with a mIoU of 69% on the real testing data. The augmented simulated data thus missed the benchmark achieved by the real data by only 4%, with the augmented sampling-based alternative performing significantly worse with a remaining delta of 9%. Relatively, the simulation-based augmented dataset outperforms the sampling-based version by 8.3%.

TABLE 6. Results for Experiment 2 on cooling plant dataset: Class-wise F_1 -score for training on datasets generated by sampling (SAM+) and simulation (SIM+) augmented by 12% real data, homogeneous real laser scan data (REAL) for reference; Δ_{SAM} , Δ_{SIM} indicate the absolute changes in comparison to Experiment 1.

	SAM+	Δ_{SAM}	SIM+	Δ_{SIM}	REAL
wall	0.97	+0.25	0.98	+0.05	0.97
floor	0.99	+0.02	0.99	+0.06	0.99
ceiling	0.99	+0.97	0.99	+0.90	1.00
beam	0.96	+0.29	0.96	+0.19	0.96
railing	0.00	-	0.00	-	0.00
pipe fitting	0.88	+0.35	0.90	+0.33	0.92
pipe accessories	0.77	+0.31	0.81	+0.26	0.89
ventilation duct	0.00	-	0.66	+0.66	0.63
cable routing	0.94	+0.46	0.95	+0.46	0.95
bracing	0.93	+0.93	0.93	+0.86	0.93
tank	0.07	+0.03	0.27	-0.01	0.32
clutter	0.65	+0.65	0.64	+0.64	0.82
equipment	0.91	+0.55	0.89	+0.47	0.91
noise	0.00	-	0.14	+0.14	0.24

designed model of the facility, those objects are not included in those point clouds.

We extended our experiment to the factory hall point cloud to investigate the potential such a hybrid dataset has in the context of a different facility. For this, a small part of the factory hall dataset was designated as augmentation data; the entire rest of the point cloud was designated as testing data. Just like in the first implementation of the experiment, the model was trained on a dataset of synthetic cooling plant data augmented with real data from the cooling plant. Subsequently, the synthetic cooling plant training data was augmented with the augmentation part of the factory hall dataset; training the model was repeated with this hybrid dataset. Finally, both of these models trained on hybrid point cloud datasets were used to perform inference on the factory hall test set; all results of this are collected in **TABLE 7**.

TABLE 7. Results for Experiment 2 on the factory hall dataset: Macro mean class Intersection over Union (mIoU) and mean class F1-score (mcF1) for varying setups, training on synthetic datasets generated by means of sampling (SAM) and simulation (SIM) augmented with real laser scanning data from cooling plant dataset (+), and factory hall dataset (*); Δ_{SAM} , Δ_{SIM} indicate the absolute changes in comparison to Experiment 1.

	SAM+	Δ_{SAM}	SIM+	Δ_{SIM}	SAM*	Δ_{SAM}	SIM*	Δ_{SIM}
mIoU	0.23	+0.08	0.26	+0.04	0.35	+0.20	0.37	+0.15
mcF1	0.29	+0.09	0.34	+0.06	0.42	+0.22	0.45	+0.17

The results show similar characteristics as in experiment 1, as the model performance on the factory hall dataset is below the results for the cooling plant. By augmentation of the synthetic cooling plant training data with real data from the cooling plant dataset, the performance increase on the factory hall dataset ranges between 5% and 10%, and the absolute added value of simulated data over sampling-based data is clear with 3% and 5%. Noticeably, the model trained on real training data from the cooling plant performs similarly to the model trained on the augmented, synthetic cooling plant dataset (cf. **TABLE 3**).

Even at this small size, the augmentation set from the factory hall leads to the best performance. The difference in performance between sampling- and simulation-based data is persistent at 2% - 3%.

4.5 Comparison with related experiments

The closest related study is the one introduced by J. W. Ma et al. (2020). In contrast to the mentioned paper's method for data generation, the sampling technique used in this paper is limited to the surface of the objects instead of the full volume of the object, which brings the presented method significantly closer to an actual laser scan that is limited to object surfaces. Furthermore, points are sampled using a random distribution instead of a grid for the presented sampling-based approach; the simulation-based version introduces TLS-specific properties such as realistic placement-dependent occlusions, range-dependent resolution, and precision. In their further studies regarding point densities, J. W. Ma et al. (2020) seem to have used random down-sampling, which lifts the spatial restriction of the grid to a certain degree.

Despite the limitations in comparability regarding dataset and sampling technique, this study shows that the simulated synthetic dataset carries significantly more information than the already improved sampling-based alternative while saving storage space. This was shown in the results of the presented experiments, where the dataset generated by simulation has shown increased value for training a neural network to perform well on real, unseen data, in comparison to sampling-based synthetic data, both for homogeneous synthetic as well as synthetic data augmented with real laser scan points.

5 DISCUSSION

Synthetic point clouds generated through simulation depict reality better than sampling-based alternatives. Specifically in the industrial domain, where LiDAR is the predominant acquisition technique and complicated geometries often lead to strong occlusions, it makes sense to apply data generation methods that result in data that is as close as possible to real data. The experiments presented in this paper investigate the value that synthetic point cloud data created using two alternative methods have as training data for neural networks for semantic segmentation.

In the two introduced experiments, diverging goals are followed. While Experiment 1 aims to clarify the added value of generating more realistic synthetic point cloud data

for training neural networks for point cloud semantic segmentation, Experiment 2 aims to match the performance of a neural network trained on real laser scanning data with a minor trade-off by adding a small subset of manually prepared ground truth data.

As both of the experiments are successful, they are substantial and show that, without pre-trained networks or large-scale labeled data, new domains can be tackled with regard to point cloud semantic segmentation, given that 3D models of comparable scenes are available.

While sampling-based methods are a valid option, employing simulation-based approaches leads to fewer overall points in the point cloud training set, saving storage space and significantly improving overall performance. In the presented method of data generation, the only necessary manual intervention remains in preparing the CAD model for simulation by assigning object classes according to the chosen class structure. This step is necessary if the available 3D model does not possess instance-level semantics – which was the case in the presented experiment. This step can be fully automated by parsing and filtering object semantics in a complete and correct BIM or 3D CAD model with full semantic information on the instance level.

Depending on the use case, the overall results of Experiment 1 might be insufficient for robust further processing due to weak results for certain classes. With the results of Experiment 2, it could be shown that with limited manual intervention, synthetic data leads to robust results that provide a good trade-off between expensively annotated real-world scan data and the cheap solution of sampling-based synthetic data.

The presented approach comes with several limitations. As the experiments have shown, the quality of the 3D model is decisive for the value of the synthetic data generated on its basis. Poor representation of actual geometric features directly influenced the generalization power of our method and led to poor prediction performance for affected classes and neighboring objects belonging to other classes.

The experiments showcase the application of the proposed method for generating synthetic point cloud data only on one facility. Furthermore, compared to other ground truth point cloud datasets for semantic segmentation, the presented three datasets representing the

cooling plant are relatively small. In general, this is a shortcoming for the training of a neural network.

To evaluate the potential of such a baseline of training data to generalize and perform inference on other datasets, a second point cloud dataset of real laser scan data was added to the experiment. The results of these extended experiments on the second dataset are in line with the findings on the first dataset. It could be shown that, with very limited additional annotation effort, the baseline training data could be extended to be useful for a significantly different dataset in terms of purpose, layout, size, and included object classes.

The comparability of the presented method is inherently limited due to the absence of established benchmarks and public data. For example, access to their models and sampled data would be required to evaluate the added value of the presented method in comparison to the work using a manually remodeled part of S3DIS and volumetric grid-based sampling (J. W. Ma et al., 2020). To this date, there is a lack of publicly available ground truth datasets showing full industrial scenes that could be used as a starting point for training their own models and benchmarking developed approaches.

6 CONCLUSION AND OUTLOOK

This paper proposes to integrate realistic synthetic ground truth data into a workflow for point cloud semantic segmentation for the industrial domain, where the absence of publicly available ground truth datasets prevents the implementation of standard approaches with readily available annotated real ground truth data for training. At the same time, the industrial domain poses specific challenges, mainly regarding geometry, resulting occlusions, and specific classes. The presented work shows that realistic synthetic data is helpful for semantic segmentation. Furthermore, compared with data generated in a sampling-based method, the synthetic data created using laser scanning simulation shows a substantial performance increase. Creating such synthetic data requires no manual effort, given suitable 3D models are available- thus, they can be generated quickly with complete, error-less class annotations. Furthermore, as they depict the laser scan in terms of occlusions and surface coverage, in comparison with a full, sampling-based

approach, around half of the overall point cloud size is sufficient for reaching the same mean surface density as the real point cloud dataset. Combined with a small amount of real laser scan data, synthetic datasets can produce results close to the presented benchmark achieved using purely real scanning data. Thus, applying scan simulations provides significant effort-saving potential in further processing steps in Scan-to-BIM.

While it has been shown that the approach yields promising results for the industrial domain, extending similar investigations with comparable parameters to a wider variety of applications, such as infrastructure or conventional office indoor spaces, would allow for more general statements about the value and limitations of this approach. Semantic segmentation, while arguably the most valuable approach for semantic enrichment in laser scan point clouds for further processing and model reconstruction currently, is inherently limited to object classes and could be enhanced significantly by instance segmentation. Another interesting continuation of this research is identified in an extension of the presented synthetic training data generation method with this instance aspect, addressing this remaining gap. Furthermore, an investigation of the applicability for different domains with specific requirements and classes seems interesting, along with the impact of model quality on the process.

ACKNOWLEDGMENTS

This work was conducted within the scope of a research project funded by AUDI AG, Ingolstadt, Germany. Further support was received in the form of hardware from NVIDIA through their Applied Research Accelerator program. The support of both companies is gratefully acknowledged.

REFERENCES

- Agapaki, E. (2020). *Automated Object Segmentation in Existing Industrial Facilities*. <https://doi.org/10.17863/CAM.52102>
- Agapaki, E., & Brilakis, I. (2020). CLOI-NET: Class segmentation of industrial facilities' point cloud datasets. *Advanced Engineering Informatics*, 45. <https://doi.org/10.1016/j.aei.2020.101121>
- Agapaki, E., Miatt, G., & Brilakis, I. (2018). Prioritizing object types for modelling existing industrial facilities. *Automation in Construction*, 96, 211–223. <https://doi.org/10.1016/j.autcon.2018.09.011>
- Amazon Web Services. (2021). *Data Labeling - Amazon SageMaker Ground Truth - AWS*. <https://aws.amazon.com/sagemaker/groundtruth/>,
- Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., & Savarese, S. (2016). 3D semantic parsing of large-scale indoor spaces Supplementary. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016- Decem*, 1534–1543. <https://doi.org/10.1109/CVPR.2016.170>
- Balado, J., Díaz-Vilariño, L., Arias, P., & Lorenzo, H. (2019). Point clouds for direct pedestrian pathfinding in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 148, 184–196. <https://doi.org/10.1016/j.isprsjprs.2019.01.004>
- Blender Online Community. (2021). *Blender - a 3D modelling and rendering package*. <http://www.blender.org>
- Borrmann, A., König, M., Koch, C., & Beetz, J. (2018). Building Information Modeling: Why? What? How? In *Borrmann A., König M., Koch C., Beetz J. (eds) Building Information Modeling*. Springer, Cham. Springer. <https://doi.org/10.1007/978-3-319-92862-3>
- Bosché, F., Ahmed, M., Turkan, Y., Haas, C. T., & Haas, R. (2015). The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction*, 49, 201–213. <https://doi.org/10.1016/j.autcon.2014.05.014>
- Cazorla, R., Poinel, L., Papadakis, P., & Buche, C. (2021). Bottleneck Identification to Semantic Segmentation of Industrial 3D Point Cloud Scene via Deep Learning. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)*, 4877–4878.
- Chern, W. C., Nguyen, T. V., Asari, V. K., & Kim, H. (2023). Impact of loss functions on semantic segmentation in far-field monitoring. *Computer-Aided Civil and Infrastructure Engineering*, 38(3), 372–390. <https://doi.org/10.1111/mice.12832>
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. *Eurographics Italian Chapter Conference*.
- CloudCompare. (2021). *CloudCompare (version 2.11.3) [GPL Software]*. <http://cloudcompare.org/>
- Corsini, M., Cignoni, P., & Scopigno, R. (2012). Efficient and Flexible Sampling with Blue Noise Properties of Triangular Meshes. *IEEE Transactions on Visualization and Computer Graphics*, 18(6), 1–12.

- Croce, V., Caroti, G., Luca, L. De, Jacquot, K., Piemonte, A., & Véron, P. (2021). From the semantic point cloud to heritage-building information modeling: A semiautomatic approach exploiting machine learning. *Remote Sensing*, 13(3), 1–34. <https://doi.org/10.3390/rs13030461>
- Czerniawski, T., Nahangi, M., Haas, C., & Walbridge, S. (2016). Pipe spool recognition in cluttered point clouds using a curvature-based shape descriptor. *Automation in Construction*, 71(Part 2), 346–358. <https://doi.org/10.1016/j.autcon.2016.08.011>
- Dai, A., Chang, A. X., Savva, M., Halber, M., Funkhouser, T., & Nießner, M. (2017). ScanNet: Richly-annotated 3D reconstructions of indoor scenes. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 2432–2443. <https://doi.org/10.1109/CVPR.2017.261>
- Deschaud, J.-E. (2021). *KITTI-CARLA: a KITTI-like dataset generated by CARLA Simulator*. 1–2. <http://arxiv.org/abs/2109.00892>
- Deschaud, J.-E., Duque, D., Richa, J. P., Velasco-Forero, S., Marcotegui, B., & Goulette, F. (2021). Paris-carla-3d: A real and synthetic outdoor point cloud dataset for challenging tasks in 3d mapping. *Remote Sensing*, 13(22). <https://doi.org/10.3390/rs13224713>
- Dosovitskiy, A., Ros, G., Codevilla, F., López, A., & Koltun, V. (2017). CARLA: An Open Urban Driving Simulator. *Proceedings of the 1st Annual Conference on Robot Learning*, 1–16.
- Esmoris, A. M., Vilariño, D. L., Arango, D. F., Varela-García, F. A., Cabaleiro, J. C., & Rivera, F. F. (2023). Characterizing zebra crossing zones using LiDAR data. *Computer-Aided Civil and Infrastructure Engineering*, 38(13), 1767–1788. <https://doi.org/10.1111/mice.12968>
- Frias, E., Pinto, J., Sousa, R., Lorenzo, H., & Díaz-Vilariño, L. (2022). Exploiting BIM Objects for Synthetic Data Generation toward Indoor Point Cloud Classification Using Deep Learning. *Journal of Computing in Civil Engineering*, 36(6). [https://doi.org/10.1061/\(asce\)cp.1943-5487.0001039](https://doi.org/10.1061/(asce)cp.1943-5487.0001039)
- Fumarola, M., & Poelman, R. (2011). Generating virtual environments of real world facilities: Discussing four different approaches. *Automation in Construction*, 20(3), 263–269. <https://doi.org/10.1016/j.autcon.2010.08.004>
- Gao, B., Pan, Y., Li, C., Geng, S., & Zhao, H. (2020). *Are We Hungry for 3D LiDAR Data for Semantic Segmentation?* 1–26. <http://arxiv.org/abs/2006.04307>
- Geiger, A., Lenz, P., Stiller, C., & Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research*, *October*, 1–6.
- Gonzalez Stefanelli, M. J., Rajesh Jain, A., Kamal Jalui, S., & Agapaki, E. (2022, July). Synthetic Point Cloud Generation for Class Segmentation Applications. *Proceedings of the 2022 European Conference on Computing in Construction*.
- Gregor, M., Medvecký, S., Matuszek, J., & Stefanik, A. (2009). Digital Factory. *Journal of Automation Mobile Robotics and Intelligent Systems*, 3, 123–132.
- Griffiths, D., & Boehm, J. (2019). *SynthCity: A large scale synthetic point cloud*. 1–6. <http://arxiv.org/abs/1907.04758>
- Gschwandtner, M., Kwitt, R., Uhl, A., & Pree, W. (2011). BlenSor: Blender sensor simulation toolbox. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6939 LNCS(PART 2), 199–208. https://doi.org/10.1007/978-3-642-24031-7_20
- Guo, M. H., Cai, J. X., Liu, Z. N., Mu, T. J., Martin, R. R., & Hu, S. M. (2021). PCT: Point cloud transformer. *Computational Visual Media*, 7(2), 187–199. <https://doi.org/10.1007/s41095-021-0229-5>
- Hammond, R. S., Nawari, N. O., & Walters, B. (2014). BIM in Sustainable Design: Strategies for Retrofitting/Renovation. *Computing in Civil and Building Engineering*, 1969–1977. <https://doi.org/doi:10.1061/9780784413616.244>
- Hong, Y., Park, S., Kim, H., & Kim, H. (2021). Synthetic data generation using building information models. *Automation in Construction*, 130(July), 103871. <https://doi.org/10.1016/j.autcon.2021.103871>
- Huang, J., Yang, X., Zhou, F., Li, X., Zhou, B., Lu, S., Ivashov, S., Giannakis, I., Kong, F., & Slob, E. (2023). A deep learning framework based on improved self-supervised learning for ground-penetrating radar tunnel lining inspection. *Computer-Aided Civil and Infrastructure Engineering*. <https://doi.org/10.1111/mice.13042>
- Hullo, J. F., & Thibault, G. (2014). Scaling up close-range surveys, a challenge for the generalization of as-built data in industrial applications. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(5), 293–299. <https://doi.org/10.5194/isprsarchives-XL-5-293-2014>
- Hullo, J.-F., Thibault, G., Boucheny, C., Dory, F., & Mas, A. (2015). Multi-Sensor As-Built Models of Complex Industrial Architectures. *Remote Sensing*, 7(12), 16339–16362. <https://doi.org/10.3390/rs71215827>
- ISO. (2018). *Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema* (Vol. 2018).
- Li, F., Li, H., & Kim, M. K. (2022). Registration-free point cloud generation technique using rotating mirrors. *Computer-Aided Civil and Infrastructure Engineering*, 37(2), 204–226. <https://doi.org/10.1111/mice.12707>
- Lu, R., & Brilakis, I. (2019). Digital twinning of existing reinforced concrete bridges from labelled point clusters. *Automation in Construction*, 105. <https://doi.org/10.1016/j.autcon.2019.102837>

- Luhmann, T., Chizhova, M., Gorkovchuk, D., Popovas, D., Gorkovchuk, J., & Hess, M. (2022). DEVELOPMENT OF TERRESTRIAL LASER SCANNING SIMULATOR. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 46(2/W1-2022), 329–334. <https://doi.org/10.5194/isprs-archives-XLVI-2-W1-2022-329-2022>
- Ma, J. W., Czerniawski, T., & Leite, F. (2020). Semantic segmentation of point clouds of building interiors with deep learning: Augmenting training datasets with synthetic BIM-based point clouds. *Automation in Construction*, 113(March), 103144. <https://doi.org/10.1016/j.autcon.2020.103144>
- Ma, L., Sacks, R., Kattel, U., & Bloch, T. (2018). 3D Object Classification Using Geometric Features and Pairwise Relationships. *Computer-Aided Civil and Infrastructure Engineering*, 33(2), 152–164. <https://doi.org/10.1111/mice.12336>
- Macher, H., Landes, T., & Grussenmeyer, P. (2017). From point clouds to building information models: 3D semi-automatic reconstruction of indoors of existing buildings. *Applied Sciences (Switzerland)*, 7(10), 1–30. <https://doi.org/10.3390/app7101030>
- Maturana, D., & Scherer, S. (2015). VoxNet: A 3D Convolutional Neural Network for real-time object recognition. *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 922–928. <https://doi.org/10.1109/IROS.2015.7353481>
- Mirzaei, K., Arashpour, M., Asadi, E., Masoumi, H., & Li, H. (2022). Automatic generation of structural geometric digital twins from point clouds. *Scientific Reports*, 12(1), 22321. <https://doi.org/10.1038/s41598-022-26307-7>
- Morbidoni, C., Pierdicca, R., Paolanti, M., Quattrini, R., & Mammoli, R. (2020). Learning from synthetic point cloud data for historical buildings semantic segmentation. *Journal on Computing and Cultural Heritage*, 13(4). <https://doi.org/10.1145/3409262>
- Neumann, M. (2020). *{Blender2Helios} - Github Repository*. <https://github.com/neumicha/Blender2Helios>
- Noichl, F., Braun, A., & Borrmann, A. (2021). “BIM-to-Scan” for Scan-to-BIM: Generating Realistic Synthetic Ground Truth Point Clouds based on Industrial 3D Models. *Proceedings of the 2021 European Conference on Computing in Construction*, 2, 164–172. <https://doi.org/10.35490/ec3.2021.166>
- Ochmann, S., Vock, R., Wessel, R., & Klein, R. (2016). Automatic reconstruction of parametric building models from indoor point clouds. *Computers and Graphics (Pergamon)*, 54, 94–103. <https://doi.org/10.1016/j.cag.2015.07.008>
- Oh, B. K., Kim, K. J., Kim, Y., Park, H. S., & Adeli, H. (2017). Evolutionary learning based sustainable strain sensing model for structural health monitoring of high-rise buildings. *Applied Soft Computing Journal*, 58, 576–585. <https://doi.org/10.1016/j.asoc.2017.05.029>
- Papers With Code. (2021). *Papers With Code: State of the Art: Semantic Segmentation on S3DIS*. <https://paperswithcode.com/sota/semantic-segmentation-on-s3dis>
- Park, H. S., Lee, H. M., Adeli, H., & Lee, I. (2007). A new approach for health monitoring of structures: Terrestrial laser scanning. *Computer-Aided Civil and Infrastructure Engineering*, 22(1), 19–30. <https://doi.org/10.1111/j.1467-8667.2006.00466.x>
- Park, S. W., Park, H. S., Kim, J. H., & Adeli, H. (2015). 3D displacement measurement model for health monitoring of structures using a motion capture system. *Measurement: Journal of the International Measurement Confederation*, 59, 352–362. <https://doi.org/10.1016/j.measurement.2014.09.063>
- Pärm, E. A., Edwards, D. J., & Sing, M. C. P. (2017). The building information modelling trajectory in facilities management: A review. In *Automation in Construction* (Vol. 75, pp. 45–55). Elsevier B.V. <https://doi.org/10.1016/j.autcon.2016.12.003>
- Perez-Perez, Y., Golparvar-Fard, M., & El-Rayes, K. (2021). Scan2BIM-NET: Deep Learning Method for Segmentation of Point Clouds for Scan-to-BIM. *Journal of Construction Engineering and Management*, 147(9), 1–14. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0002132](https://doi.org/10.1061/(ASCE)CO.1943-7862.0002132)
- Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). PointNet: Deep learning on point sets for 3D classification and segmentation. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 77–85. <https://doi.org/10.1109/CVPR.2017.16>
- Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems, 2017-Decem*, 5100–5109.
- Richter, S. R., Vineet, V., Roth, S., & Koltun, V. (2016). Playing for data: Ground truth from computer games. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9906 LNCS, 102–118. https://doi.org/10.1007/978-3-319-46475-6_7
- Rottensteiner, F., Sohn, G., Gerke, M., & Wegner, J. D. (2013). ISPRS Test Project on Urban Classification and 3D Building Reconstruction. *ISPRS - Commission III - Photogrammetric Computer Vision and Image Analysis Working Group III / 4 - 3D Scene Analysis*, 1–16.
- Roynard, X., Deschaud, J. E., & Goulette, F. (2018). Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. *International Journal of Robotics Research*, 37(6), 545–557. <https://doi.org/10.1177/0278364918767506>
- Sharif, M. M., Nahangi, M., Haas, C., & West, J. (2017). Automated Model-Based Finding of 3D Objects in Cluttered Construction Point Cloud Models. *Computer-Aided Civil and Infrastructure Engineering*, 32(11), 893–908. <https://doi.org/10.1111/mice.12306>

- Shellshear, E., Berlin, R., & Carlson, J. S. (2015). Maximizing Smart Factory Systems by Incrementally Updating Point Clouds. *IEEE Computer Graphics and Applications*, 35(2), 62–69. <https://doi.org/10.1109/MCG.2015.38>
- Shi, X., Xu, X., Chen, K., Cai, L., Foo, C. S., & Jia, K. (2021). *Label-Efficient Point Cloud Semantic Segmentation: An Active Learning Approach*.
- Smith, A., & Sarlo, R. (2022). Automated extraction of structural beam lines and connections from point clouds of steel buildings. *Computer-Aided Civil and Infrastructure Engineering*, 37(1), 110–125. <https://doi.org/10.1111/mice.12699>
- Soilán, M., Nóvoa, A., Sánchez-Rodríguez, A., Justo, A., & Riveiro, B. (2021). Fully automated methodology for the delineation of railway lanes and the generation of IFC alignment models using 3D point cloud data. *Automation in Construction*, 126(February). <https://doi.org/10.1016/j.autcon.2021.103684>
- Talebi, S. (2014). Exploring advantages and challenges of adaptation and implementation of BIM in project life cycle - University of Salford Institutional Repository. *2nd BIM International Conference on Challenges to Overcome*. <http://usir.salford.ac.uk/id/eprint/32275/>
- Thomas, H. (2021). *KPConv-PyTorch*. <https://github.com/HuguesTHOMAS/KPConv-PyTorch>
- Thomas, H., Qi, C. R., Deschaut, J. E., Marcotegui, B., Goulette, F., & Guibas, L. (2019). KPConv: Flexible and deformable convolution for point clouds. *Proceedings of the IEEE International Conference on Computer Vision, 2019-October*, 6410–6419. <https://doi.org/10.1109/ICCV.2019.00651>
- Tong, Z., Ma, T., Zhang, W., & Huyan, J. (2023). Evidential transformer for pavement distress segmentation. *Computer-Aided Civil and Infrastructure Engineering*. <https://doi.org/10.1111/mice.13018>
- VDA. (2009). *3D Datenaustausch in der Fabrikplanung*.
- Volk, R., Stengel, J., & Schultmann, F. (2014). Building Information Modeling (BIM) for existing buildings - Literature review and future needs. *Automation in Construction*, 38(October 2017), 109–127. <https://doi.org/10.1016/j.autcon.2013.10.023>
- Walsh, S. B., Borello, D. J., Guldur, B., & Hajjar, J. F. (2013). Data processing of point clouds for object detection for structural engineering applications. *Computer-Aided Civil and Infrastructure Engineering*, 28(7), 495–508. <https://doi.org/10.1111/mice.12016>
- Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., & Solomon, J. M. (2019). Dynamic Graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 38(5), Article 146.
- Wang, Z., Zhang, Y., Mosalam, K. M., Gao, Y., & Huang, S. L. (2022). Deep semantic segmentation for visual understanding on construction sites. *Computer-Aided Civil and Infrastructure Engineering*, 37(2), 145–162. <https://doi.org/10.1111/mice.12701>
- Wiendahl, H.-P., Reichardt, J., & Nyhuis, P. (2015). *Handbook Factory Planning and Design*. <https://doi.org/https://doi.org/10.1007/978-3-662-46391-8>
- Winiwarter, L., Pena, A. M. E., Weiser, H., Anders, K., Sanchez, J. M., Searle, M., & Höfle, B. (2021). *Virtual laser scanning with HELIOS++: A novel take on ray tracing-based simulation of topographic 3D laser scanning*. 1–31. <http://arxiv.org/abs/2101.09154>
- Wu, Y., Qin, Y., Qian, Y., Guo, F., Wang, Z., & Jia, L. (2022). Hybrid deep learning architecture for rail surface segmentation and surface defect detection. *Computer-Aided Civil and Infrastructure Engineering*, 37(2), 227–244. <https://doi.org/10.1111/mice.12710>
- Xie, Y., Tian, J., Zhu, X. X., & Member, S. (2019). *A Review of Point Cloud Semantic Segmentation*. 1–51. <http://arxiv.org/abs/1908.08854>
- Yin, C., Wang, B., Gan, V. J. L., Wang, M., & Cheng, J. C. P. (2021). Automated semantic segmentation of industrial point clouds using ResPointNet++. *Automation in Construction*, 130(July), 103874. <https://doi.org/10.1016/j.autcon.2021.103874>
- Yokoyama, H., Date, H., Kanai, S., & Takeda, H. (2013). Detection and Classification of Pole-like Objects from Mobile Laser Scanning Data of Urban Environments. *International Journal of CAD/CAM*, 13(2), 31–40. <http://147.47.106.98/ojs/index.php/ijcc/article/view/247>
- Zhai, R., Zou, J., He, Y., & Meng, L. (2022). BIM-driven data augmentation method for semantic segmentation in superpoint-based deep learning network. *Automation in Construction*, 140, 104373. <https://doi.org/10.1016/j.autcon.2022.104373>
- Zhang, J., Zhao, X., Chen, Z., & Lu, Z. (2019). A Review of Deep Learning-Based Semantic Segmentation for Point Cloud. *IEEE Access*, 7, 179118–179133. <https://doi.org/10.1109/ACCESS.2019.2958671>
- Zheng, Y., Gao, Y., Lu, S., & Mosalam, K. M. (2022). Multistage semisupervised active learning framework for crack identification, segmentation, and measurement of bridges. *Computer-Aided Civil and Infrastructure Engineering*, 37(9), 1089–1108. <https://doi.org/10.1111/mice.12851>
- Zhou, Y., Huang, J., Dai, X., Luo, L., Chen, Z., & Ma, Y. (2020). *HoliCity: A City-Scale Data Platform for Learning Holistic 3D Structures*.
- Zhou, Y., Xiang, Z., Zhang, X., Wang, Y., Han, D., & Ying, C. (2022). Mechanical state inversion method for structural performance evaluation of existing suspension bridges using 3D laser scanning. *Computer-Aided Civil and Infrastructure Engineering*, 37(5), 650–665. <https://doi.org/10.1111/mice.12765>