# Prototype pipeline modelling using interval scanning point clouds

**Toa Pečur[1] · Frédéric Bosché[1] · Gabrielis Cerniauskas[1] ·
Frank Mill[1] · Andrew Sherlock[2] · Nan Yu[1]**

**Abstract** With the aid of computer aided design (CAD) and building information modelling (BIM), as-built to as-designed comparison has seen many developments in improving the workflow of manufacturing and construction tasks. Recently, evolution has been centred around automation of scene interpretation from three-dimensional (3D) scan data. The scope of this paper is to assess assemblies as the installation process progresses and inferring if arising deviations are problematic (complex task). The adequacy of utilising unorganised point clouds to compliance check are trialled with a real life down-scaled prototype model in conjunction with a synthetic dataset. This work aims to highlight areas where large rework could be avoided, in part by the detection of potential clashes of components early in the pipeline installation process. With the help of an extracted model in the form of a point cloud generated from a scanned physical model and a 3D CAD model representing the nominal geometry, an operator can be made visually aware of potential deviations and component clashes during a pipeline assembly process.

**Keywords** Point clouds · Modelling · Computer aided design (CAD) · Digital manufacturing · Interval scanning

✉ Nan Yu
nan.yu@ed.ac.uk

1 School of Engineering, The University of Edinburgh, Edinburgh EH9 3FB, Scotland, UK

2 National Manufacturing Institute Scotland, Paisley PA3 2EF, Scotland, UK

## 1 Introduction

To enable the comparison of as-design computer aided design (CAD) or building information modelling (BIM) models with the as-built state of construction, a three-dimensional (3D) scan of the environment is required, usually from a scanner which outputs point clouds. Information on the comparison can then be used for a variety of applications from maintenance of structures [1] to collision detection when installing or disassembling components [2]. Comparing deviations from scan to CAD/BIM have been explored by numerous researchers such as Bosché et al. [3, 4] and Anil et al. [5]. Comparing deviations to models whose parts change significantly from the CAD/BIM are complex tasks as the final location of the parts could be in a significantly different pose (differing in translation and/or rotation) as well as hindered by occlusion effects from the scanning process. Utilising the local information becomes less robust with larger deviations not documented in CAD/BIM and requires manual operator intervention. However, not all deviations are critical, implying that deviations should be further examined to assess the impact on the sub-assemblies in place.

In this study, the focus is on keeping an updated version of the as-is state of a piping assembly and sub-assemblies during the installation process utilising point clouds. The study aims to integrate quality control during the installation of the piping sub-assemblies and assembly without the explicit need of separate quality assurance stations for prefabrication. Sub-assemblies are not routinely compared with neighbouring sub-assemblies in quality control, albeit rather to identify deviations or defects within the current sub-assembly [6]. This study aims to reduce rework by giving the operator the ability and flexibility to detect and remediate major deviations earlier. Indirectly, this reduces

the time for installing the components at the final location and lowers the tolerances experienced when installing the assemblies on site. This provides greater oversight from the current progress of the installation and enables documentation for maintenance purposes (as-is CAD/BIM is created in the process). The point cloud processing occurs with an automated robust workflow during installation, reducing the manual processing required by the operator. The recognition of the valves and other components is undertaken with a combination of robust registration and deep learning(DL) methods. The DL model is trained using synthetic data that are generated using an automated process and effectively imitates real-life scans.

Section 2 provides an overview of the state of the art actively applied when undertaking similar pipeline tasks. Section 3 presents the method followed through this study, where experimental setup was showcased in Sect. 4 and 3D scanning detailed within Sect. 5. Results are presented in Sect. 6 and further discussed in Sect. 7.

## 2 Background

### 2.1 Pipeline workflows

Using 3D point clouds obtained from scanners, there have been works for construction comparisons to CAD/BIM such as Bosché and Haas [7]. The context of pipe and valve components has been explored by various authors. Maalek et al. [8] explored automated pipe and flanges extraction using a framework of robust circle fitting techniques. Kim et al. [9] investigated pipe and elbow automated recognition using curvature information with primitive classification with simple and S-shaped elbow. Sharif et al. [10] employed local feature of point pairs for object recognition and isolation adapted from the framework based on Papazov and Burschka [11] for mechanical, electrical, plumbing (MEP) components. As-built pipe network reconstruction has been considered by Kawashima et al. [12]. Pipe spool recognition work was outlined by Czerniawski et al. [13] for automation when

extracting spools in an as-built scene using an extraction algorithm based on local curvature estimation, clustering and a bag-of-features matching. Inspection of pipe systems was validated by Nguyen and Choi [14] using normal-based region growing and efficient random sample consensus in shipbuilding compartments. Deviation comparisons to as-is was detailed by Anil et al. [15] using point clouds of a plant facility which concluded with a time reduction of identifying errors by 40% when using conventional measurement methods. Safa et al. [16] showed a measurement process when undertaking piping fabrication with certain areas marked as critical points on the pipeline as well as quality assurance stations for measuring the components when completed. In the case where a previous 3D model is unavailable, scan-to-BIM methods are used such as the reconstruction of pipeline by Cheng et al. [17].

Less common is utilising 3D point clouds with in situ measurements to generate an updated digital model (associated 3D model) as the operator works through the piping installation, where deviations are varied (small and large). An example not in the scope of pipeline installation, but
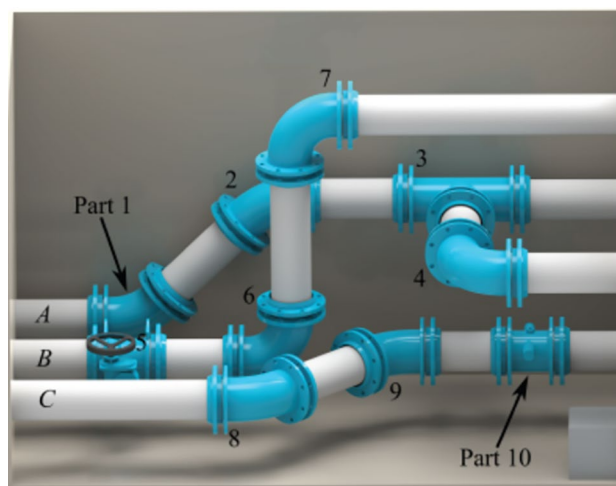


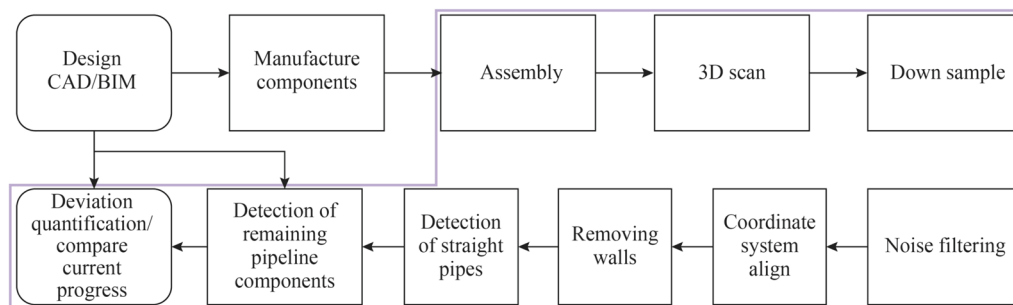**Fig. 2** CAD digital rendering of 3PM with starting and ending part numbering



**Fig. 1** Workflow flowchart (task repetition highlighted in purple)

related to it, is seen by Kim et al. [18] who conducted quality inspection for full-scale precast concrete elements from laser scanning and BIM. Another distant example is seen in a brief study with pipe installation by Fraga-Lamas et al. [19] within shipbuilding using multiple sensors rather than point clouds.

## 2.2 Point cloud classification using DL

To further decrease the burden of manual operation, automated classification labelling operations have been actively developed to ease the task of semantic interpretation. Greater focus has been concentrated on the advancement of 3D DL architectures available for point cloud classification as they have lagged behind their 2D counterpart models. Partly due

to the complexity of processing 3D information with the added difficulty from dealing with unorganised data with no coherent order. PointNet [20], Pointnet++ [21], DGCNN [22], and PCT [23] are examples of point cloud DL models attempting to bridge that gap.

## 2.3 Data for DL model training

With the majority of DL models currently requiring supervised learning for training, synthetic generated benchmarks are vital for an understanding of DL model capabilities and are readily used when comparing different models (such as dataset ModelNet40 [24] and PartNet [25]). In contrast, in comparison to real life scans, it is more difficult to directly classify objects accurately employing synthetic data for
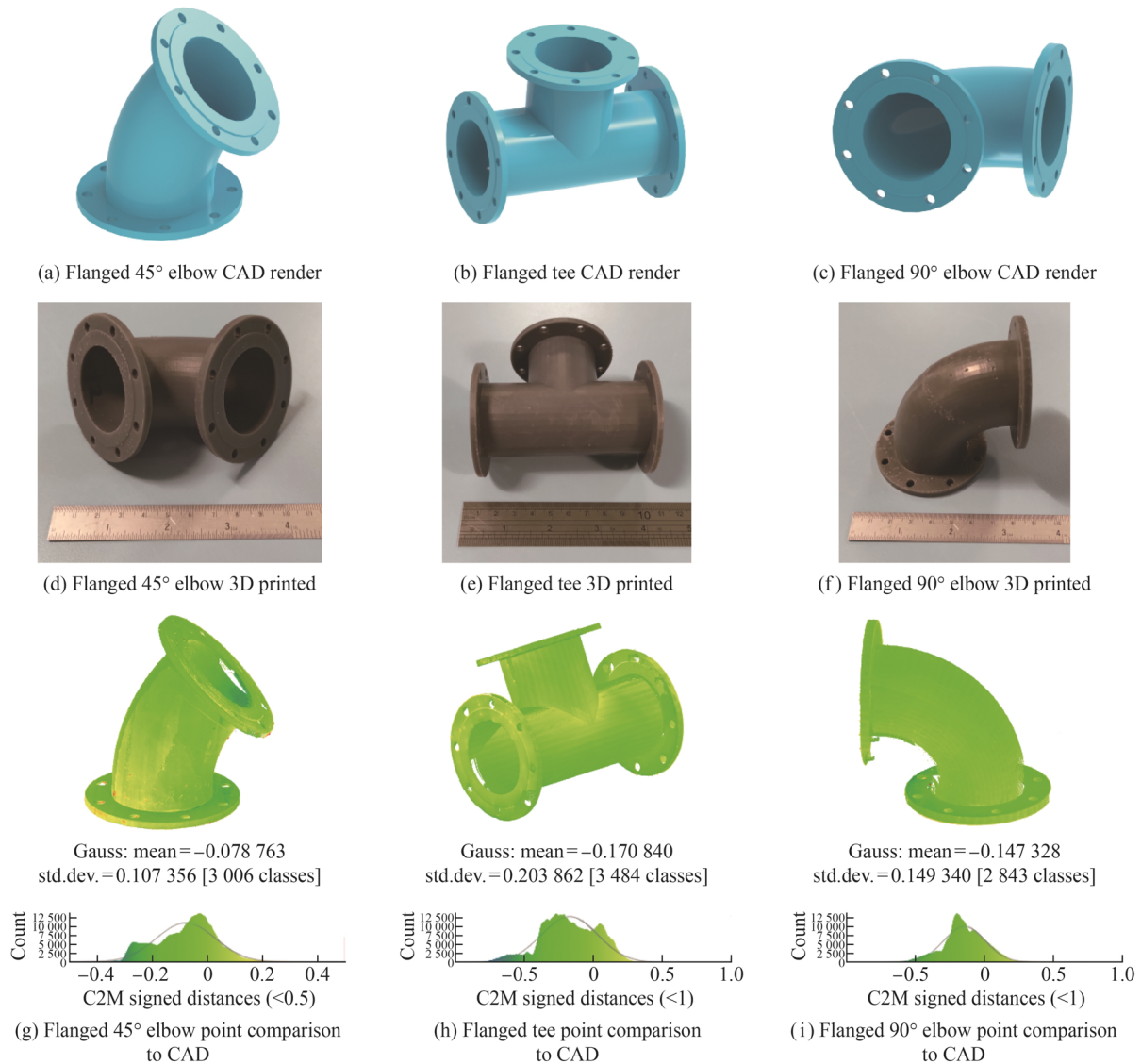


(a) Flanged 45° elbow CAD render

(b) Flanged tee CAD render

(c) Flanged 90° elbow CAD render

(d) Flanged 45° elbow 3D printed

(e) Flanged tee 3D printed

(f) Flanged 90° elbow 3D printed

Gauss: mean=−0.078 763 std.dev.=0.107 356 [3 006 classes]

Gauss: mean=−0.170 840 std.dev.=0.203 862 [3 484 classes]

Gauss: mean=−0.147 328 std.dev.=0.149 340 [2 843 classes]

(g) Flanged 45° elbow point comparison to CAD

(h) Flanged tee point comparison to CAD

(i) Flanged 90° elbow point comparison to CAD

**Fig. 3** Prototype downscaled DN200 PN10 parts

training purposes using homogeneous discretisation over entire surfaces. As by Uy et al. [26], real world scanning data still prove to be complex when undertaking classification for DL tasks. Their study showed, that while the accuracy of CAD like dataset was reaching near perfection, training a model to classify a real-world object dataset was still a very challenging task which led them to the creation of a custom real-world object dataset. This can be explained as DL models are primarily being compared and benchmarked on clean, non-corrupted CAD datasets.

Real-life pipework and valve datasets such as from Yeo et al. [27] are useful and have been incorporated into DL tasks, for example in Ref. [28] during classification of pipework components. Nevertheless, limited pipework and valve point cloud datasets (both synthetic and real life) are publicly available encompassing different types and varieties of design standards which could be reliably trained on. In addition to the limitations of the variety of valve and elbow datasets, datasets with a large number of point clouds for training a variety of occasions and conditions are limited as well.

Datasets for semantic segmentation require per point labelling of data which is a time-consuming process requiring human intervention. There have been work from authors such as Yin et al. [29] which have created four industry areas of point clouds for semantic segmentation, and required the combined use of human operators and DL to categorise pipeline scenes for semantic segmentation purposes [30]. Recently, Korus et al. [31] discussed using DynamoPC-Sim for synthetic point cloud creation to reduce the cost and time of manual data creation using Revit and visual programming.

Evidently, there is a need for a workflow expediting creation of large training-agnostic datasets, which minimizes the laborious task of categorisation and includes real-life aspects of the scanning procedure arising from pipeline assembly.
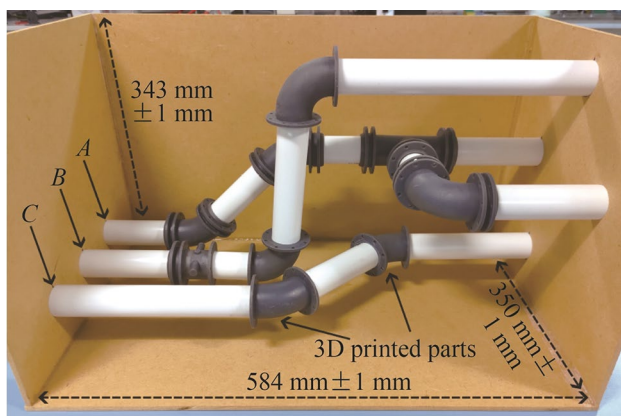


**Fig. 4** 3PM setup physical dimensions with sub-assemblies "*A*", "*B*", and "*C*" indicated

## 3 Process and method overview

We propose a method, outlined in Fig. 1, where, as input, the CAD/BIM designer creates an assembly of a model which then gets submitted to manufacturing. This model undergoes assembly with 3D interval scanning as the operator progresses through the steps until completion. The interval 3D point clouds obtained from the scanner are down sampled for quicker processing turnaround times. Noise filtering removes spurious points due to the 3D scanner machine characteristics. Next, undertaking coordinate system alignment facilitates direct comparisons of the progression to the CAD assembly model. Segmentation of homogeneous surface geometries is used to detect straight pipe segments. This is followed by clustering of the remaining points to be classified into seven groups by a deep neural network. The final stage is comparing the classified objects with the CAD/BIM objects using robust registration to obtain the as-is pose of each individual CAD part within the assembly (see Sect. 6.4). This process is repeated until the sub-assemblies and the whole assembly are completed. Following each iteration of this process, an updated CAD model is overlaid with the original position of the sub-assemblies for the operator to visualize.

## 4 Experimental model setup

The steps of the proposed pipeline are illustrated (and validated) using a downscaled physical prototype pipeline installation model (3PM). The 3PM was created to showcase the numerous possibilities for assembly deviations during installation from design intent. Figure 2 shows a rendering of the CAD model of that prototype.

The pipeline (containing three different sub-assemblies) was envisioned to be built in a modular capacity and has an arbitrary valve/elbow configuration, where compartments could be created off-site and connected at a later setting in the final location. Flange couplings are considered as a core strategy for prefabrication in this work as weld and butt fusion joints are not deemed to be easily detachable for ease of use in such a scenario, as commented by Li et al. [32] in a study on pipe prefabrication. The smaller scale of the model (1 : 5.5̇) allows for easier trialling of potential solutions without the inconvenience of moving larger and heavier parts. The parts were numerically numbered starting from the left hand side of the CAD sub-assembly "*A*" finishing on the right hand side of sub-assembly "*C*".

The CAD models were selected for a size of DN200 and were downscaled by a factor of 5.5̇ to retain characteristics of geometrical components obtained when scanning in real life with a light detection and ranging (LiDAR) scanner. The models were 3D printed using stereolithography
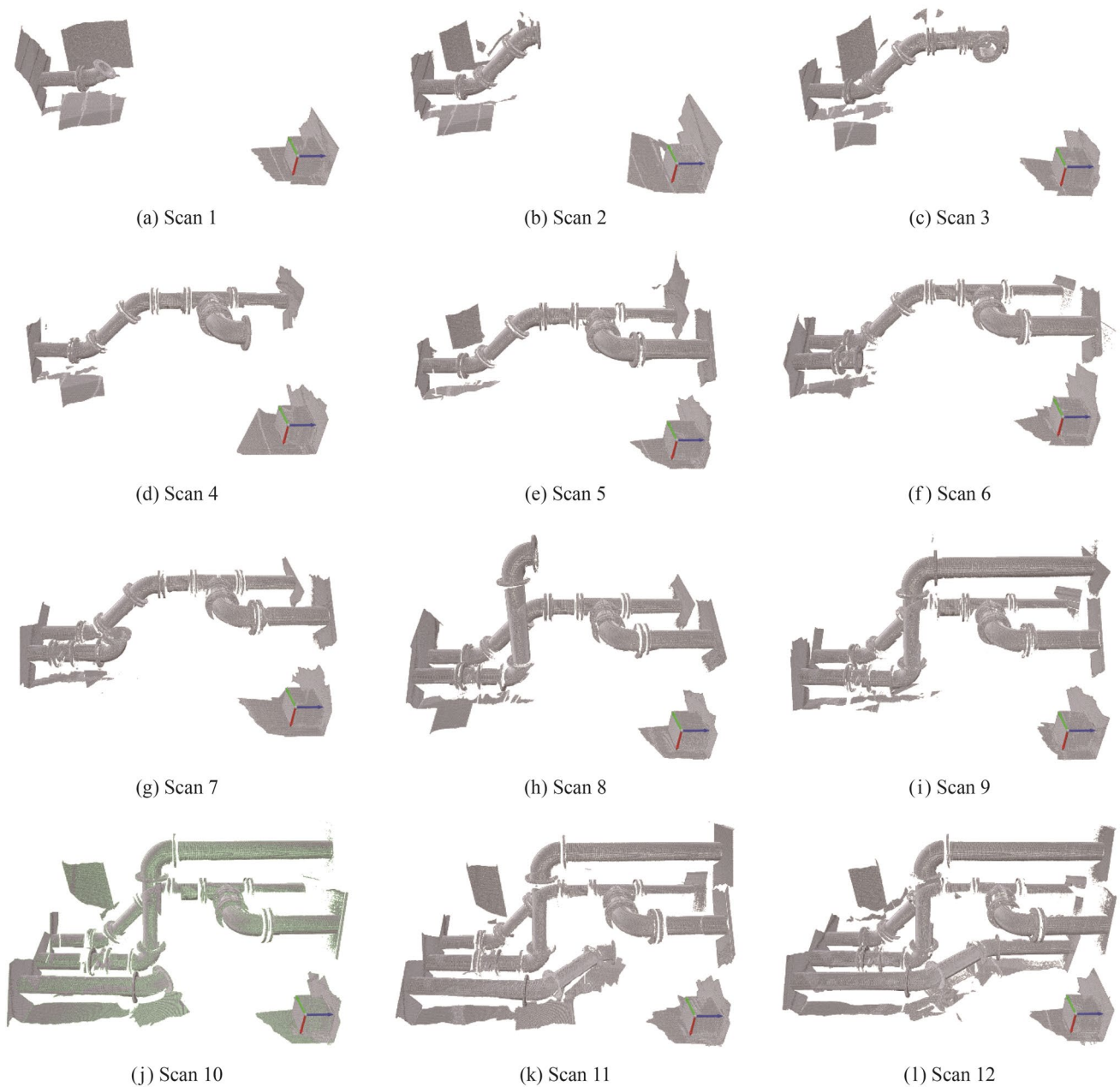
(a) Scan 1

(b) Scan 2

(c) Scan 3

(d) Scan 4

(e) Scan 5

(f) Scan 6

(g) Scan 7

(h) Scan 8

(i) Scan 9

(j) Scan 10

(k) Scan 11

(l) Scan 12

**Fig. 5** Scan interval point clouds using Faro LLP on 3PM

(SLA) grey resin from designs of a manufacturer of valve and piping components (see Figs. 3d, e and f). SLA printing is used to ensure accurate dimensionality containing intricate details which have been verified by individual part scanning and comparing to the dimensions of the CAD model seen in Figs. 3a b, c. Orientation of the parts on the bed was adjusted to minimise the structural support that would be seen when scanning. Curing was used to enhance the rigidity of the parts for handling when installing the components. The printed parts are a tight press fit for the polyvinyl chloride (PVC) pipe of outer diameter

36 mm as no adjustments were made to compensate for the slight shrinking experienced, as seen in Figs. 3g, h, i.

The 3PM was created to contain three distinct pipelines ("A", "B", "C") labelled in Fig. 4 similar to the CAD model designed (see Fig. 2). Sub-assembly "A" was first installed followed by "B" and "C", as this allowed optimum scanning while building the pipeline assembly. The installed parts have both translation and rotation deviations from their CAD counterparts, as well as valves in incorrect sub-assembly positions (part 9, check valve) and missing in sub-assemblies (part 5, butterfly valve). A mixture of different changes
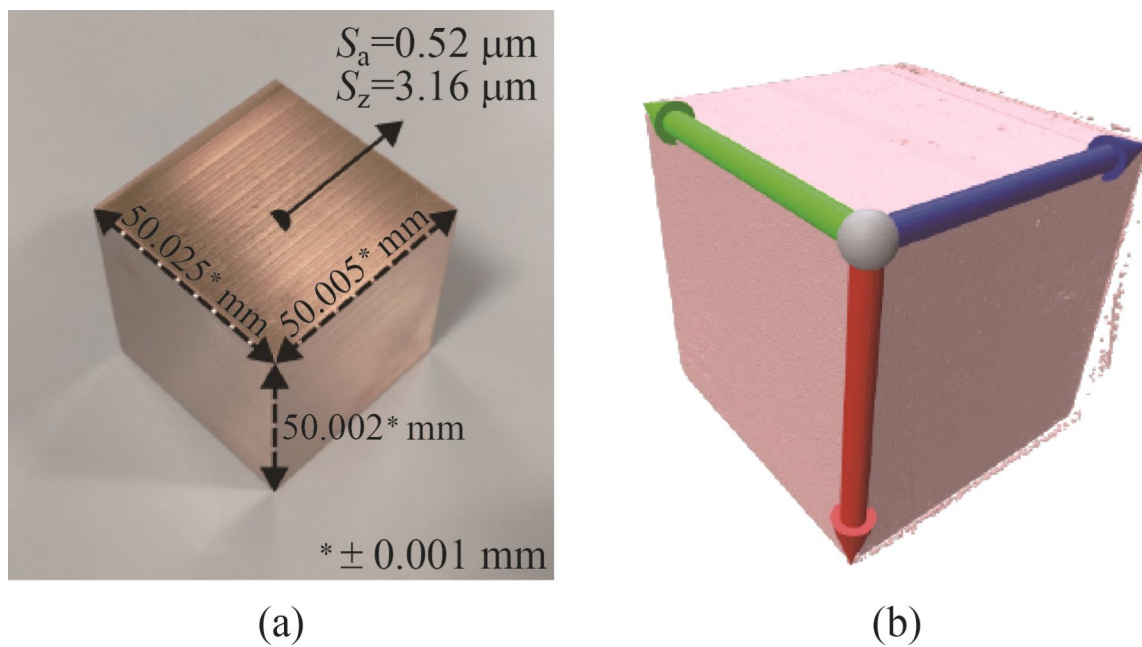
**Fig. 6** Invar 36 cube "50 mm × 50 mm × 50 mm" peripheral milled surface side used for global origin alignment when **a** 3D scanning, **b** global system alignment using three planes
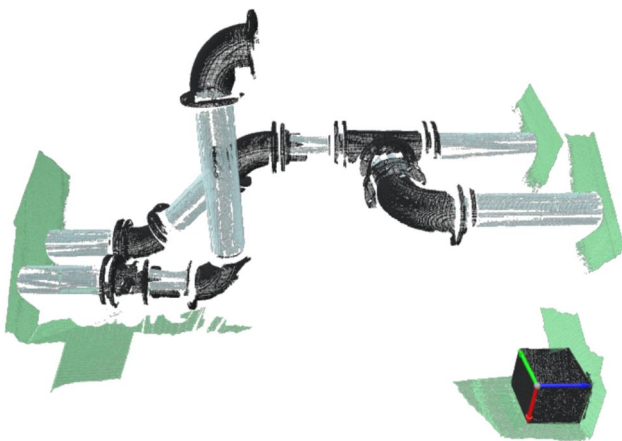


**Fig. 7** Scan 8 showing plane RANSAC segmentation (light green), cylinder RANSAC segmentation (turquoise)

from the digital ground truth model were created to trial the accuracy of the workflow in the ability to extract CAD components.

## 5 Model scanning

The assembly and 3D models were scanned using a non-contact Faro laser line probe (LLP) connected to a portable Faro ScanArm computer measuring machine (CMM) which was error compensated beforehand with an error of

0.023 8 mm. No scanner filtering settings were used during the scanning. The 3PM was scanned at intervals when additional valve or flange components were installed and for a final pass when the whole sub-assembly was completed. This allows ample time for the operator to know when a pipeline would likely become problematic and could remedy the installation before major rework would be required (at the completion stage). Each scan of the scene is treated as a different setup seen in Fig. 5, where the CMM LLP scanner was not moved during each individual scan. During the installation of the sub-assemblies, deviations due to the lack of supporting structures are also present in the scene. This requires a final scan pass of each sub-assembly once both pipe ends meet the hole cut-outs.

During LLP scanning, the world origin is centred around the base of the Faro ScanArm. To undertake comparisons to the assembly model each scan also includes the Invar 36 cube (see Fig. 6) within the scene. This assists (for the down-scaled models) the coordinate system alignment using the three planes method from the planes extracted on the Invar 36 cube which is fixed in place. Intersection of three planes is considered as the world origin point satisfying Cartesian Eq. (1), where the intersection of two planes is used as a vector for the $z$ direction (blue) and the subsequent $x$ and $y$ direction are best oriented to the intersection of the other two planes [33]. The normal plane vector is represented by $\boldsymbol{n} = \langle \hat{a}, \hat{b}, \hat{c} \rangle$ for each plane.

**Table 1** Comparison of virtual training dataset using various DN values

| | All flanged cross | Flanged tee | Butterfly valve | Check valve | Cube | Flanged 45° elbow | Flanged 90° elbow |
|---|---|---|---|---|---|---|---|
| | DN100 PN10/16 | DN50 PN10/16 | DN150 PN16 | DN200 PN16 | Six sides | DN50 PN10/16 | DN50 PN10/16 |
| Sampled points | | | | | | | |
| Virtual LiDAR | | | | | | | |
| | DN300 PN16 | DN250 PN10 | DN200 PN16 | DN450 PN16 | Five sides | DN200 PN10 | DN100 PN10/16 |
| Sampled points | | | | | | | |
| Virtual LiDAR | | | | | | | |

$$\begin{cases} \hat{a}_1 x + \hat{b}_1 y + \hat{c}_1 z + \hat{d}_1 = 0, \\ \hat{a}_2 x + \hat{b}_2 y + \hat{c}_2 z + \hat{d}_2 = 0, \\ \hat{a}_3 x + \hat{b}_3 y + \hat{c}_3 z + \hat{d}_3 = 0. \end{cases} \tag{1}$$

For the downscaled workflow, this strategy allows for a smaller area coverage when undertaking alignment which benefits from less scanning required by the operator. In a full-size construction environment, this registration method could be adjusted by instead utilising the surrounding walls and floor that would have previously undergone quality control, and thus the as-is positions and shapes would have already been corrected in the CAD/BIM model. Different alignment strategies could also be adapted when undertaking precision alignment of large scientific instruments that require module modular assembly. This allows for more flexibility and planning when working while not at a final location of the site.

## 6 Comparisons to CAD

After the scanning is completed by the operator, raw point clouds from the LLP scanner are downsampled for quicker processing speeds. Voxel-based downsampling with the same parameter was used for all twelve scans to facilitate capturing the underlying surface.

### 6.1 Prepossessing (removing walls)

Efficient random sample consensus (RANSAC) [34] is used to detect the large planar regions from the sides (walls and floors) of the 3PM. For this, we used the same parameters throughout the twelve different scans. Figure 7 illustrates the result of this process with planar segment detections shown in green.
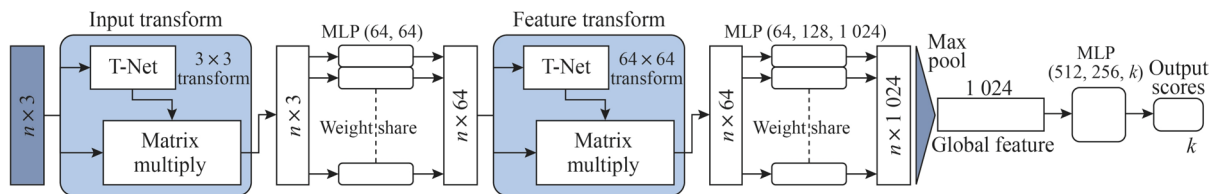


**Fig. 8** PointNet classification network
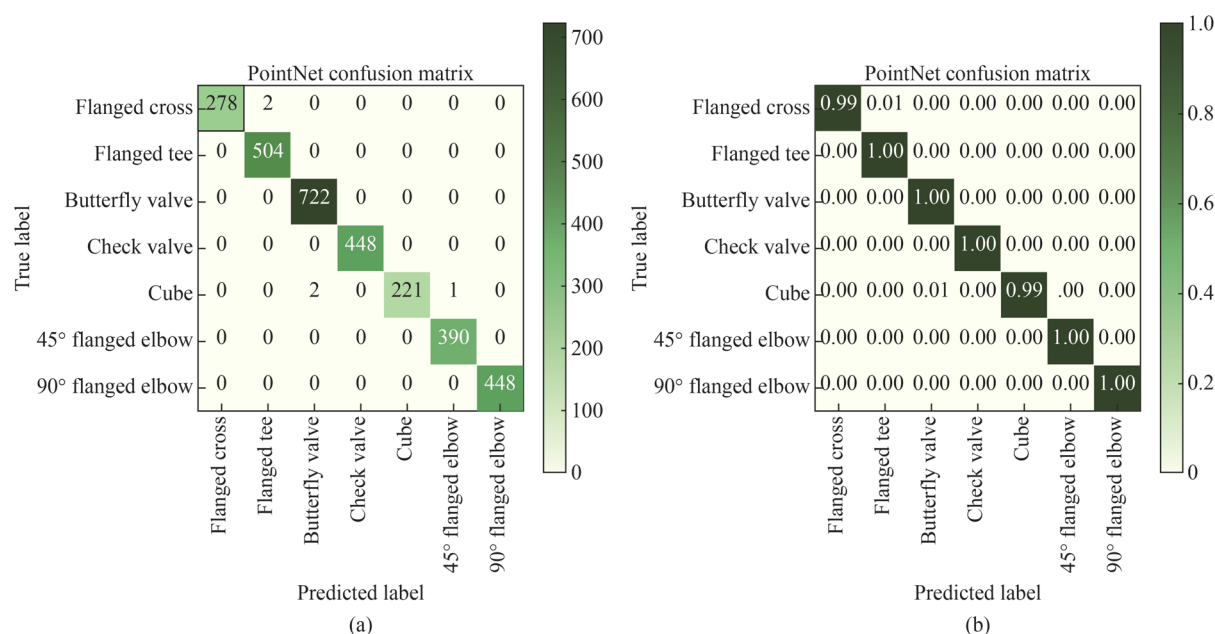
**Fig. 9** PointNet validation data confusion matrix **a** virtual LiDAR scanner dataset, **b** the normalized matrix
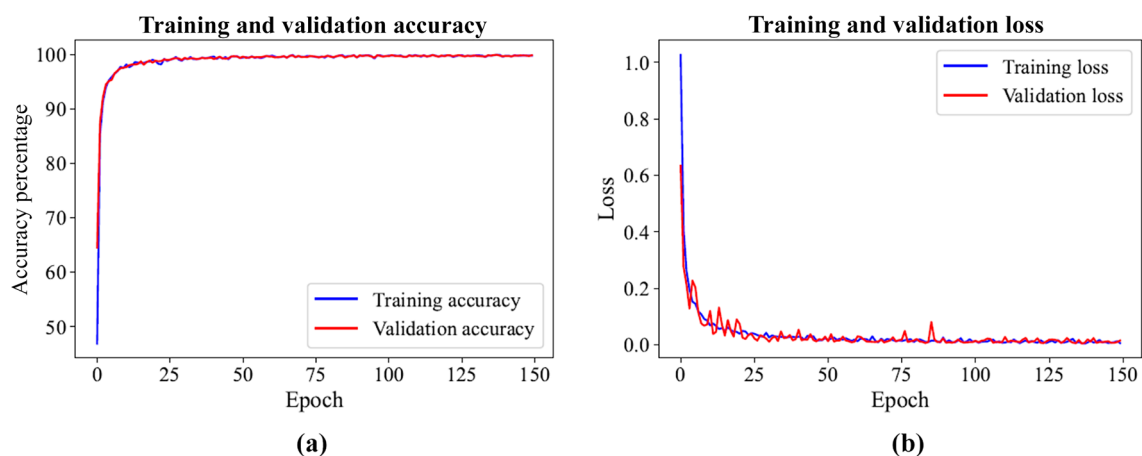


**Fig. 10** PointNet **a** accuracy, and **b** loss during training with virtual LiDAR scanner dataset

### 6.2 Detection of straight pipes

RANSAC is then used again to detect all the straight pipe sections within the scene as cylinders. For this, we again used the same parameters throughout the twelve scans. Figure 7 illustrates the result of this process with straight pipe segment detections shown in turquoise colour.

### 6.3 Detection of remaining pipeline components

Density-based spatial clustering of applications with noise (DBSCAN) [35] is then used to cluster the remaining point clouds in corresponding groups for cylinders and

components. In the case of very small clusters that are incorrectly grouped, each smaller cluster is filtered by automatically removing point clouds with a sparse number of points. The remaining component clusters require to be labelled to their corresponding categories.

The remaining clusters are finally classified using deep learning model. The model is developed using a supervised training method using 3D CAD models of common pipeline components. Commonly, models of CAD components have been grouped using seven common categories: elbow, pump, valve, etc. Removing the requirement from labelling the data and enabling the use of existing name grouping structures already devised. This resulted in seven groups

of classification, with a total of 347 CAD models. Distinct complementary CAD models have been obtained from various valve and fixture manufacturers, relevantly designed to the standards EN 545 (joints) [36], EN 593 (butterfly valves) [37], and EN 1092-2 (flange sizes) [38]. Straightforwardly, additional CAD models could be incorporated in the future to cater different industry standards. The pipe sizes used for training were mixed to include dimensions from DN10 to DN1500 and pressure ranges from PN10 to PN16 some of which could be seen in Table 1.

The CAD models in the devised dataset are split into 293 for training and 54 for validation. During the training process only virtually generated data from the CAD parts are used, with real-life individual scans not incorporated in the data. This allows for the creation of a customised dataset with numerous standardised parts for pipework and valves which are not required to be extensively modelled within assemblies for labelling. To alleviate a few of the shortcomings from synthetic data, the CAD dataset is scanned by a virtual scanner which mimics the machine characteristics of the Riegl vz400i LiDAR terrestrial laser scanner (TLS) utilising Helios++ library from Winiwarter et al. [39]. The dataset created has similar features to Wang et al. [40], as well as Huang and You [41], but with the addition of machine noise characteristics with each scan (combined with the effects of self occlusion). The synthetic dataset is LiDAR realistic with beam divergence from the scanner and full waveform representation of the 3D point clouds generated through ray tracing. Additionally, density and noise variations are baked into the point cloud dataset arising from the scanner. This workflow enables multiple different geometrical dimensions of flange and valve components allowing for greater flexibility when interpreting scenes. The synthetically created large unorganised point cloud dataset contains $1.63 \times 10^4$ point clouds for training and $3 \times 10^3$ point clouds for validating.



(a)                              (b)

**Fig. 11** **a** Scan 1, 45° flanged elbow (part 1) CAD registered to segmented point cloud, **b** scan 5, 90° flanged elbow (part 4) CAD registered to segmented point cloud

These point clouds are then prepared for batch training with a point cloud size of 4 096 which was achieved using farthest point sampling.

We have chosen PointNet as the architecture used for the classification as this model accounts for permutation invariance of point clouds, using a multilayer perceptron (MLP) layer taking the points from three dimensions to 64 dimensions, where the detailed architecture could be seen in Fig. 8. Mapping the points into 64 dimensions by having the MLP share their weights allows for an arbitrary number of points.

Each (unlabelled) point cloud cluster remaining from the previous step (straight pipe detection) is passed through the classification network of PointNet [20] with default values which was implemented with the previously custom trained CAD dataset of valve and flanges. Figure 9 reports the confused matrix obtained from the validation data.

The performance of DL models is directly impacted by data corruption levels (occlusion, noise, density, etc.) as explored by Sun et al. [42]. In their study there was a triple performance degradation observed with the same synthetic dataset when using point cloud corruptions on the data. The highest corruptions were observed when data were in LiDAR-type form (i.e., only covering a portion of the object surface) as opposed to data covering a greater surface. PointNet was reported to contain superior characteristics in standard training (without any data augmentation methods) as it produced the lowest error rates amongst other network architectures when compared to DGCNN, PCT, etc. The error was shown to be further reduced using the RSMix [43] data augmentation method.

In this study, we have increased the classification accuracy of PointNet during training (see Fig. 10) using the custom CAD dataset when compared to random sampling of point clouds. Data augmentations included object rotations and different rotations from the LiDAR scanner. The improvement is visible in Fig. A of Appendix where all 12 scans provided a substantial increase in classification accuracy when dealing with real world data.

### 6.4 Deviation quantification

The classified clusters must be compared to their as-designed 3D CAD counterparts to assess the positional deviations of the parts. To quantify this deviation, robust registration is applied to get the transformation between the point cloud cluster and the as-design 3D CAD model (mesh). Registration is done in a two-step process, where first global registration from Zhou et al. [44] is performed followed by local robust iterative closest point (ICP) registration. For the coarse alignment of the corresponding point cloud to the CAD model, fast point feature histogram (FPFH) [45] is used as the initial correspondence set $\mathcal{K}$, where for each query point $p_q$ the point feature histogram (PFH) descriptor
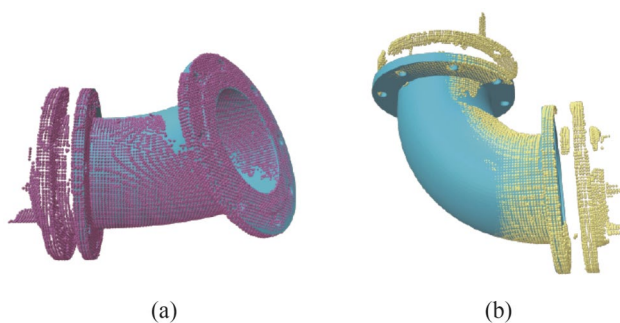
Part 1′:
$\Delta t_x = -49.2$ mm, $\Delta r_x = -1.0°$
$\Delta t_y = 25.4$ mm, $\Delta r_y = -2.3°$
$\Delta t_z = -14.1$ mm, $\Delta r_z = -6.9°$
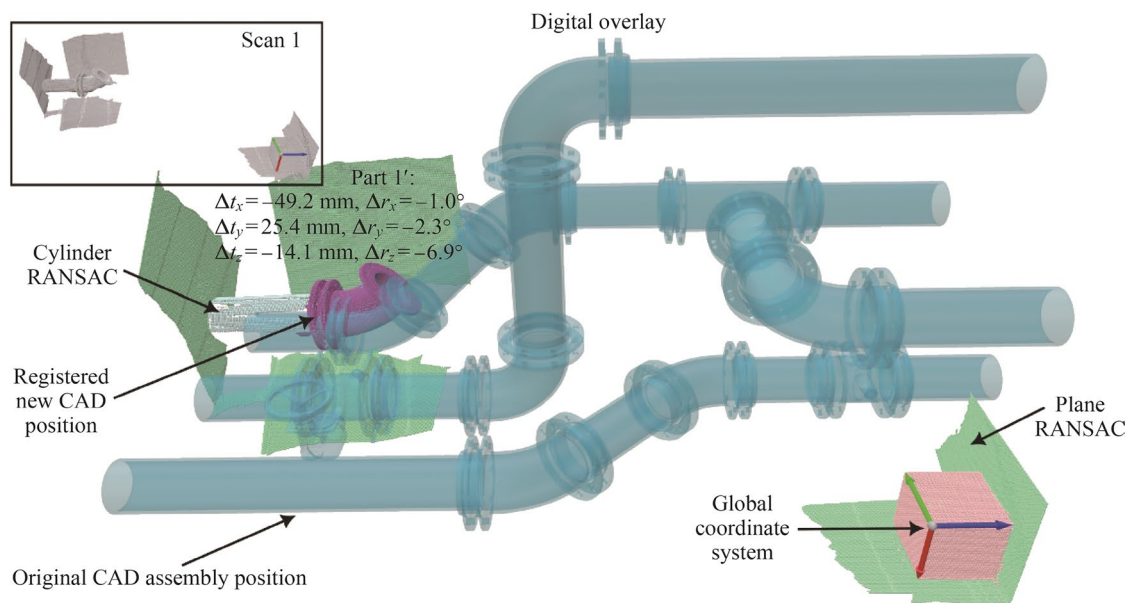
(a) Scan 1 with overlay of original CAD locations and registered new CAD locations of 45° flanged elbow (sub-assembly "A" installation)
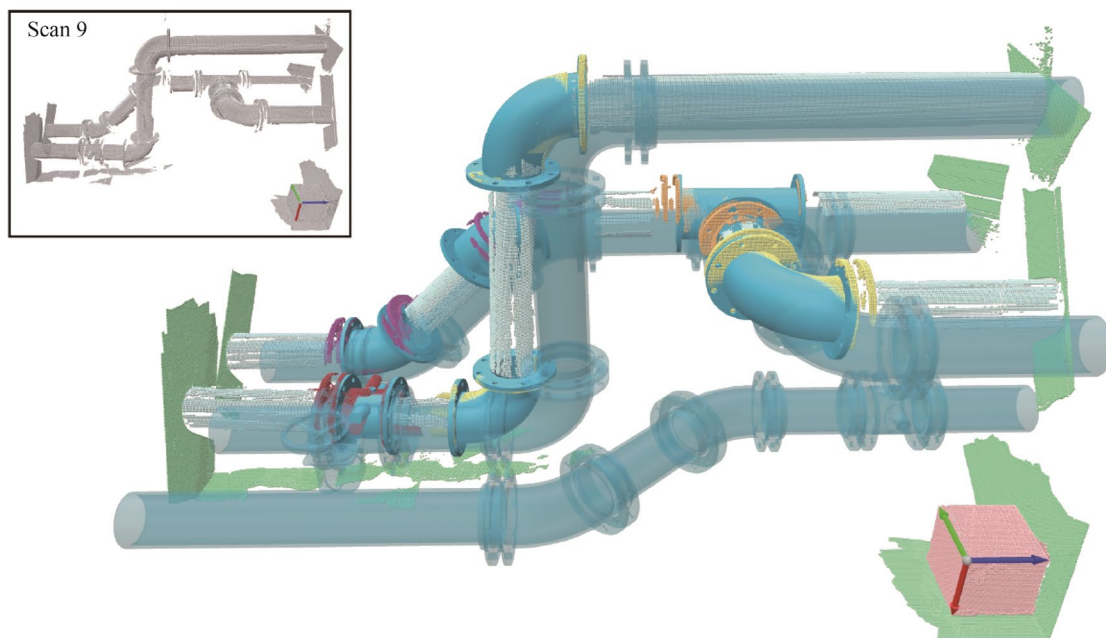


(b) Scan 9 with overlay of original CAD locations and registered new CAD locations (sub-assembly "A" and "B" installation)

Fig. 12 Digital overlay of real life scan point clouds over CAD sub-assemblies

**Fig. 13** As-designed CAD translation and rotational movement during interval scanning vs original CAD position (Part five was not found in the scene when scanning)



**Fig. 14** Scan 12 **a** 90° flanged elbow, **b** 45° flanged elbow, and **c** check valve point clouds from LLP

is calculated with the weight $w_i$ being the distance between the query point $p_q$ and a neighbor point $p_q$ as seen from Eq. (2).
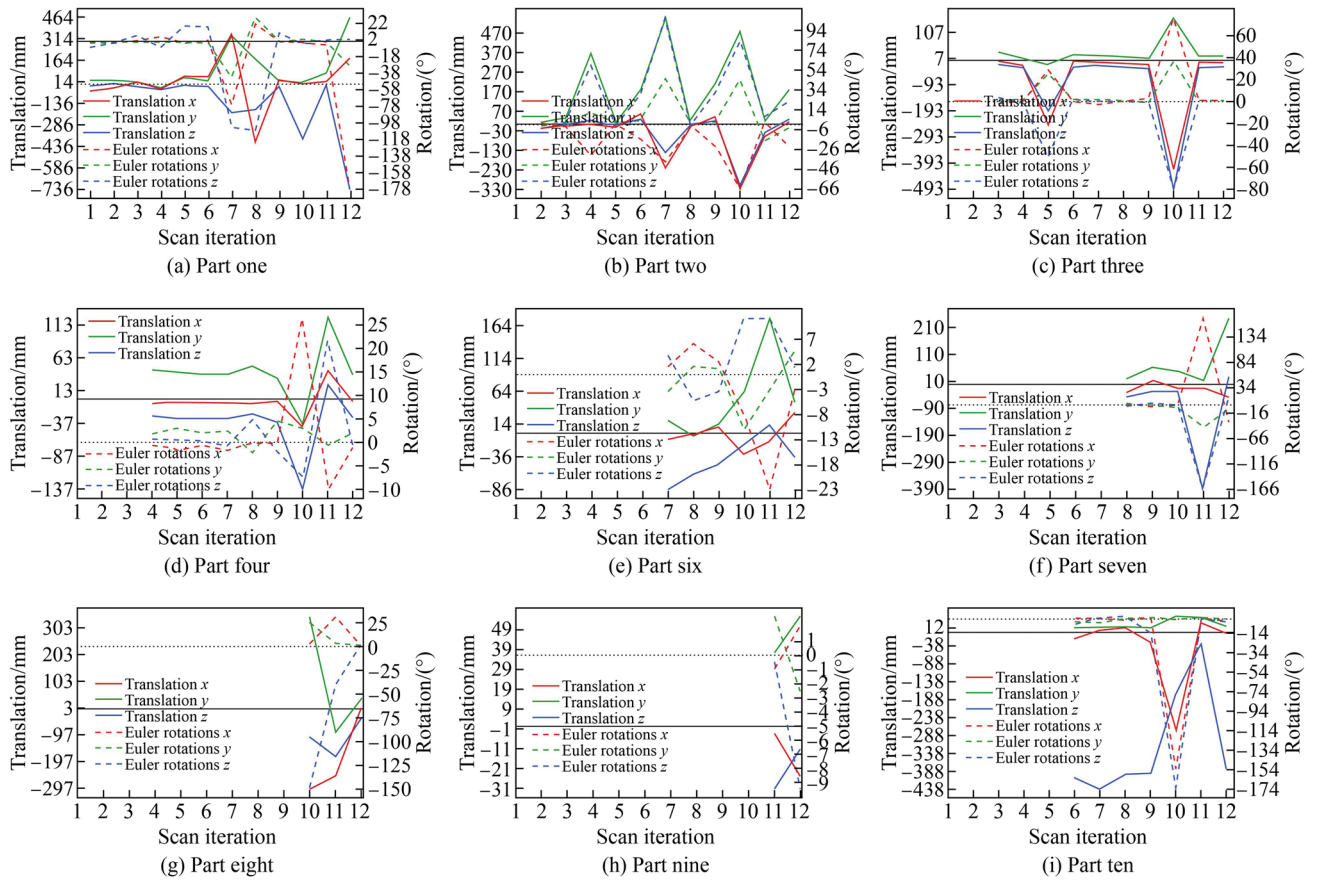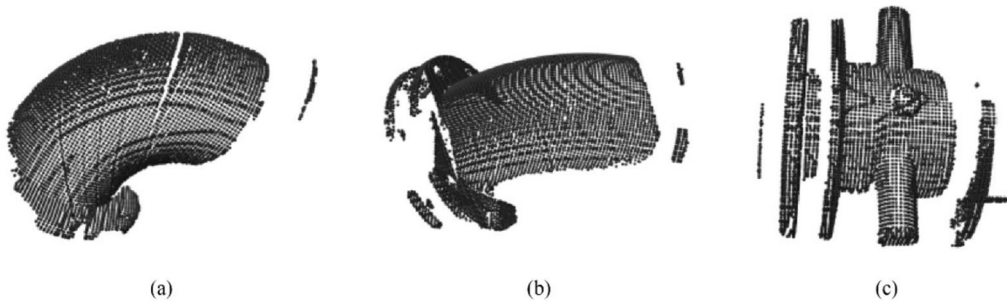
$$\text{FPFH}(p_q) = \text{PFH}(p_q)\frac{1}{k}\sum_{i=1}^{k}\left(\frac{1}{w_i}\text{PFH}(p_i)\right). \tag{2}$$

Point-to-plane ICP (see Eq. (3)) was used for fine registration of the segmented point cloud onto the CAD with a robust kernel. This is undertaken using optimization of the objective function $E(T)$, by finding the correspondence set $\mathcal{K} = \{(p,q)\}$ from the target point cloud defined as $P$, to the source point cloud defined as $Q$. Where $n_p$ is the normal of point $p$, and $\rho(r)$ is the defined as the robust kernel (Tukey was used).

$$E(T) = \sum_{(p,q)\in\mathcal{K}} \rho\big((p - T_q)n_p\big) = \sum_{i=1}^{N} \rho\big(r_i(T)\big). \tag{3}$$

Segmentation clusters can contain neighbouring welded flanges or pipe segments that might have missing information along the surface of the elbows and valves. Despite such possible errors, utilising this multi-registration workflow allows for robust registration in the case of over-segmentation with previous clusters (see Fig. 11).

This workflow is repeated for every scene scan, where the CAD locations after each scan are overlaid with the original sub-assembly positions. When the operator is installing the corresponding components, this enables them to visualize likely upcoming clashes and problematic deviations, as illustrated in Figs. 12a and b with the total 12 scans seen in Fig. B of Appendix.

## 7 Discussion

During the installation of a new flanged part, the classification and registration worked well as the occlusion was lower and neighbouring components did not create large amount of outliers. This is seen with examples in Fig. 13 where the first scan of each new component yields lower levels of registration failure. Consequently, the experimental results indicated that for better segmentation of components with lower levels of occlusion, there was a higher chance of correct classification of objects with low likelihood of registration failure. Due to the geometric nature of the parts containing similar smooth patches, it was more difficult to classify and register

45° and 90° flanged elbows than more unique components such as the flanged check valve (see Fig. 14).

The higher levels of occlusions of pipeline "A" components as the assembly progressed is attributed to:

(i) Increasingly difficult areas to reach for scanning as more pipework are physically blocking scanning access.

(ii) Less attention paid to sub-assemblies that have previously been set in place and scanned with lower surface area coverage. It is presumed the objects have not been significantly displaced. As seen in part one (from sub-assembly "A") which has increased registration failure after the sixth scene scanning (after sub-assembly "A" was completed).

Different scanning setups could be accommodated to streamline the capture of point clouds using an increased amount of either stationary 3D scanners or smaller portable 3D handheld scanners. Stationary cameras would require preregistration of the point cloud scans to reduce the high occlusion present in the scene that was experienced. Similarly, smaller portable 3D handheld scanners could be operated to improve movement in a confined space for increased flexibility when dealing with harder to reach areas.

During the training we have solely used the CAD models with no combinations of assembly geometries. Whereas, during the inference there was presence of both welded flanges near the CAD models (on either/one side) and others where welded flanges were left out on purpose which have not been used during training. Using robust registration assisted with coarse segmentation showed that classification and registration could be undertaken correctly as seen previously in Fig. 11.

Presently, PointNet has varying results on the classification, and further testing would be needed to ascertain the underlying characteristic relating to occlusion and outlier effects with similar geometrically models for classification. The current workflow could be further improved with finer semantic segmentation results of pipes from other components before passing the clusters through the classification networks. During this study we have only considered *xyz* coordinates and their CAD components. Colour, intensity, and other attributes have not been taken into account which would further aid the workflow. Using colour information could be further

combined with augmented reality/virtual reality enabling exploration using headsets for visualisations in a future study. The scanner used for the testing of this workflow in real life is an LLP, which differs from the characteristics of the Riegl scanner that was used in virtual data creation, despite also being based on laser point capturing. Current results indicate, even with the different characteristics, there is still a benefit in utilising data generation this way. The limitations are with longer slender pipes that RANSAC would have difficulty distinguishing from piping components such was the case with the flanged tee.

Finally, to provide more intuitive instructions to the operator, auto pipe (re)routing could be implemented such as in Park and Storch [46] where different objective functions for collective properties could be specified (number of elbows, spatial clearance for maintenance, maximization of support sharing with installation, compliance for vibration of nearby pipes, etc.) or similar applications [47–49].

## 8 Conclusions

Using the outlined workflow comparisons of interval-based scans to CAD enables more foresight to be given to the operator during assembling. An automated approach was devised in the detection of various pipeline elements. DL was used in the classification of non-wall and non-straight-pipe clusters with a model trained using a synthetic dataset generated from detailed geometrical CAD models. During the training multiple, different standard pipe sizes were used. Using the workflow explored, an updated assembly model of the current work as well as documentation for future maintenance can be created. In addition, costly rework can be avoided. In the case where the sub-assemblies arrive damaged due to shipping, on-site repairs could be undertaken with more certainty with the final compliance/specification using this workflow. In situ measurements have been useful in manufacturing and recently with the wider availability of both LiDAR and handheld scanners, more deployment into installation processes could be implemented.

## Appendix

See Figs. A and B.

**Fig. A** Average classification accuracy using PointNet testing real life data with virtual LiDAR and common random sampling training comparison
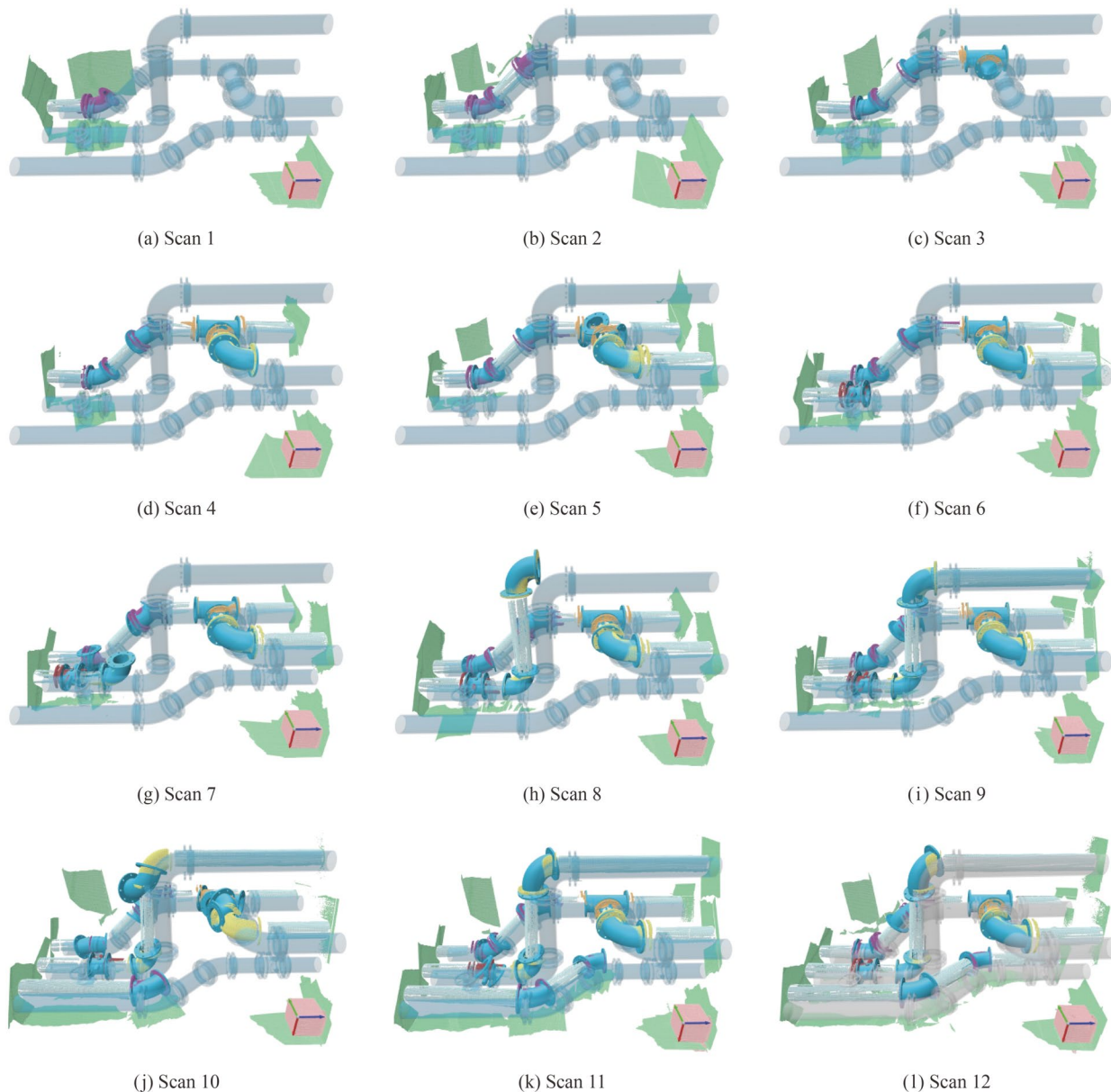
**Fig. B** As-designed CAD translation overlayed on original CAD positions with point clouds during 3PM installation

## References

1. Wang Q, Kim MK (2019) Applications of 3D point cloud data in the construction industry: a fifteen-year review from 2004 to 2018. Adv Eng Inform 39:306–319
2. Kardos C, Váncza J (2018) Application of generic CAD models for supporting feature based assembly process planning. Proc CIRP 67:446–451
3. Bosché F (2010) Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. Adv Eng Inform 24(1):107–118

4. Bosché F, Guillemet A, Turkan Y et al (2014) Tracking the built status of MEP works: assessing the value of a scan-vs-BIM system. J Comput Civ Eng 28(4):05014004. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000343

5. Anil EB, Tang P, Akinci B et al (2011) Assessment of the quality of as-is building information models generated from point clouds using deviation analysis. In: Beraldin JA, Cheok GS, McCarthy MB et al (eds) Three-dimensional imaging, interaction, and measurement, vol. 7864. International Society for Optics and Photonics, SPIE, p 78640F. https://doi.org/10.1117/12.876554

6. Nahangi M, Haas CT (2014) Automated 3D compliance checking in pipe spool fabrication. Adv Eng Inform 28(4):360–369

7. Bosché F, Haas C (2008) Automated retrieval of 3D CAD model objects in construction range images. Autom Constr 17(4):499–512

8. Maalek R, Lichti DD, Walker R et al (2019) Extraction of pipes and flanges from point clouds for automated verification of pre-fabricated modules in oil and gas refinery projects. Autom Constr 103:150–167

9. Kim Y, Nguyen CHP, Choi Y (2020) Automatic pipe and elbow recognition from three-dimensional point cloud model of industrial plant piping system using convolutional neural network-based primitive classification. Autom Constr 116:103236. https://doi.org/10.1016/j.autcon.2020.103236

10. Sharif M, Nahangi M, Haas C et al (2017) Automated model-based finding of 3D objects in cluttered construction point cloud models. Comput Aided Civ Infrastruct Eng 32(11):893–908

11. Papazov C, Burschka D (2011) An efficient RANSAC for 3D object recognition in noisy and occluded scenes. In: Kimmel R, Klette R, Sugimoto A (eds) Computer vision—ACCV 2010. Springer, Berlin Heidelberg, pp 135–148

12. Kawashima K, Kanai S, Date H (2014) As-built modeling of piping system from terrestrial laser-scanned point clouds using normal-based region growing. J Comput Des Eng 1(1):13–26

13. Czerniawski T, Nahangi M, Haas C et al (2016) Pipe spool recognition in cluttered point clouds using a curvature-based shape descriptor. Autom Constr 71:346–358

14. Nguyen CHP, Choi Y (2018) Comparison of point cloud data and 3D CAD data for on-site dimensional inspection of industrial plant piping systems. Autom Constr 91:44–52

15. Anil EB, Tang P, Akinci B et al (2013) Deviation analysis method for the assessment of the quality of the as-is building information models generated from point cloud data. Autom Constr 35:507–516

16. Safa M, Shahi A, Nahangi M et al (2015) Automating measurement process to improve quality management for piping fabrication. Structures 3:71–80

17. Cheng L, Wei Z, Sun M et al (2020) DeepPipes: learning 3D pipelines reconstruction from point clouds. Graph Models 111:101079. https://doi.org/10.1016/j.gmod.2020.101079

18. Kim MK, Wang Q, Park JW et al (2016) Automated dimensional quality assurance of full-scale precast concrete elements using laser scanning and BIM. Autom Constr 72:102–114

19. Fraga-Lamas P, Noceda-Davila D, Fernández-Caramés T et al (2016) Smart pipe system for a Shipyard 4.0. Sensors 16(12):2186. https://doi.org/10.3390/s16122186

20. Qi CR, Su H, Mo K et al (2016) PointNet: deep learning on point sets for 3D classification and segmentation. Comput Vis Pattern Recognit. https://doi.org/10.48550/arXiv.1612.00593

21. Qi CR, Yi L, Su H et al (2017) PointNet++: deep hierarchical feature learning on point sets in a metric space. Comput Vis Pattern Recognit. https://doi.org/10.48550/arXiv.1706.02413

22. Phan AV, Nguyen ML, Nguyen YLH et al (2018) DGCNN: a convolutional neural network over large-scale labeled graphs. Neural Netw 108:533–543

23. Guo M, Cai J, Liu Z et al (2021) PCT: point cloud transformer. Comput Vis Media 7(2):187–199

24. Wu Z, Song S, Khosla A et al (2014) 3D shapeNets: a deep representation for volumetric shapes. Comput Vis Pattern Recognit. https://doi.org/10.48550/arXiv.1406.5670

25. Mo K, Zhu S, Chang AX et al (2019) Partnet: a large-scale benchmark for fine-grained and hierarchical part-level 3d object understanding. In: IEEE/CVF conference on computer vision and pattern recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019

26. Uy MA, Pham QH, Hua BS et al (2019) Revisiting point cloud classification: a new benchmark dataset and classification model on real-world data. In: IEEE/CVF international conference on computer vision (ICCV), 27 October– 02 November, Seoul, Korea, pp 1588–1597. https://doi.org/10.1109/ICCV.2019.00167

27. Yeo C, Kim S, Kim H et al (2020) Deep learning applications in an industrial process plant: repository of segmented point clouds for pipework components. JMST Adv 2(1):15–24. https://doi.org/10.1007/s42791-019-00027-y

28. Yin C, Cheng JC, Wang B et al (2022) Automated classification of piping components from 3D LiDAR point clouds using SE-PseudoGrid. Autom Constr 139:104300. https://doi.org/10.1016/j.autcon.2022.104300

29. Yin C, Wang B, Gan VJL et al (2021) Automated semantic segmentation of industrial point clouds using ResPointNet++. Autom Constr 130:103874. https://doi.org/10.1016/j.autcon.2021.103874

30. Agapaki E, Brilakis I (2020) CLOI-NET: class segmentation of industrial facilities' point cloud datasets. Adv Eng Inform 45:101121. https://doi.org/10.1016/j.aei.2020.101121

31. Korus K, Czerniawski T, Salamak M (2023) Visual programming simulator for producing realistic labeled point clouds from digital infrastructure models. Autom Constr 156:105126. https://doi.org/10.1016/j.autcon.2023.105126

32. Li X, Li Z, Wu G (2017) Modular and offsite construction of piping: current barriers and route. Appl Sci 7(6):547. https://doi.org/10.3390/app7060547

33. Pečur T, Yu N, Sherlock A et al (2023) Robust coordinate system alignment using high density point clouds from laser line probe. In: Proceedings of laser metrology and machine performance XV, Edinburgh, United Kingdom, pp 103–113

34. Schnabel R, Wahl R, Klein R (2007) Efficient RANSAC for point-cloud shape detection. Comput Graph Forum 26(2):214–226

35. Ester M, Kriegel HP, Sander J et al (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: ACM SIGKDD conference on knowledge discovery and data mining, Portland, 2–4 Aug, pp 226–231

36. EN 545 (2010) Ductile iron pipes, fittings, accessories and their joints for water pipelines—requirements and test methods

37. EN 593 (2017) Industrial valves—metallic butterfly valves for general purposes

38. EN 1092-2 (1997) Flanges and their joints—circular flanges for pipes, valves, fittings and accessories, PN designated Part 2. Cast iron flanges

39. Winiwarter L, Esmorís Pena AM, Weiser H et al (2022) Virtual laser scanning with HELIOS++: a novel take on ray tracing-based simulation of topographic full-waveform 3D laser scanning. Remote Sens Environ 269:112772. https://doi.org/10.1016/j.rse.2021.112772

40. Wang H, Liu Q, Yue X et al (2021) Unsupervised point cloud pre-training via occlusion completion. In: Proceedings of the IEEE/CVF international conference on computer vision (ICCV), Montreal, 10–17 Oct, pp 9782–9792

41. Huang J, You S (2013) Detecting objects in scene point cloud: a combinational approach. In: International conference on 3D vision, IEEE, Seattle, WA, USA, pp 175–182. https://doi.org/10.1109/3DV.2013.31

42. Sun J, Zhang Q, Kailkhura B et al (2022) Benchmarking robustness of 3d point cloud recognition against common corruptions. arXiv 2201.12296. https://doi.org/10.48550/arXiv.2201.12296

43. Lee D, Lee J, Lee J et al (2021) Regularization strategy for point cloud via rigidly mixed sample. arXiv 2102.01929. https://doi.org/10.48550/arXiv.2102.01929

44. Zhou QY, Park J, Koltun V (2016) Fast global registration. In: Leibe B, Matas J, Sebe N et al (eds.) Computer vision—ECCV 2016. Lecture Notes in Computer Science, vol 9906. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-46475-6_47

45. Rusu RB, Blodow N, Beetz M (2009) Fast point feature histograms (FPFH) for 3D registration. In: IEEE international conference on robotics and automation, Kobe, 12−17 May, pp 3212–3217. https://doi.org/10.1109/ROBOT.2009.5152473

46. Park JH, Storch RL (2002) Pipe-routing algorithm development: case study of a ship engine room design. Expert Syst Appl 23(3):299–309

47. Dong Z, Bian X (2020) Ship pipe route design using improved a* algorithm and genetic algorithm. IEEE Access 8:153273–153296

48. Kim SH, Ruy WS, Jang BS (2013) The development of a practical pipe auto-routing system in a shipbuilding cad environment using network optimization. Int J Naval Archit Ocean Eng 5(3):468–477

49. Kim Y, Lee K, Nam B et al (2023) Application of reinforcement learning based on curriculum learning for the pipe auto-routing of ships. J Comput Des Eng 10(1):318–328

**Gabrielis Cerniauskas** Postdoctoral Researcher at the Institute for Materials and Processes, School of Engineering at the University of Edinburgh. His current research investigates thermoplastic liquid composite moulding and strives towards a more sustainable development of glass and carbon composites. Recently, he obtained a Ph.D. degree in computational design and development of mechanical metamaterials, with a focus on leveraging optimisation and machine learning methods for efficient design processes. Prior to this, he obtained a Master of Engineering degree in Mechanical Engineering from the University of Edinburgh. Dr. Cerniauskas' research interests lay in computational design, optimisation, machine learning, smart structures, cellular solids and digital manufacturing.

**Toa Pečur** Ph.D. candidate at the Institute for Materials and Processes at the University of Edinburgh (UoE). Having completed his Master of Engineering with Honours in Mechanical Engineering with Management in UoE, his current focus is on large scale metrology with AI digital aspects when manufacturing.
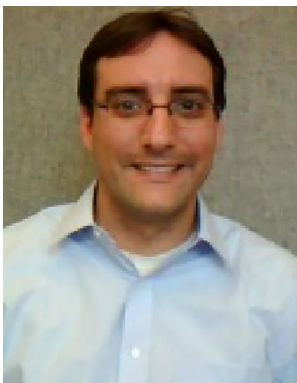
**Frédéric Bosché** Reader in Construction Informatics at the Institute for Infrastructure and Environment, the University of Edinburgh (UoE). Following a Ph.D. in Civil Engineering at the University of Waterloo (Canada), he worked for 2 years as researcher in the Computer Vision Laboratory at ETH Zurich, before becoming Assistant Professor in Construction Informatics at Heriot-Watt University. In 2019, He joined the UoE where he was first Senior Lecturer and now Reader. He teaches on construction management and informatics, and leads the Cyber Build Lab that delivers research and innovation in these areas.

**Frank Mill** Chartered engineer and Honorary Professor of Digital Design at The University of Edinburgh. He graduated with a B.Sc. (honours: first class) Technology with Industrial Studies in 1983 from Napier College and with a Ph.D. in Shape Optimisation in 2004 from the University of Edinburgh. He has worked with Redcastle Systems and ShapesSpace Ltd. on technical aspects of CAD and he is a co-founder and director of ShapeSpace Ltd. He has carried out extensive research, consultancy and teaching in CAD.

**Andrew Sherlock** Director of Data-Driven Manufacturing at the National Manufacturing Institute Scotland and Professor of Practice at the University of Strathclyde. His first degree was in Mechanical Engineer and his Ph.D. focused on novel shape optimisation techniques for aerospace components. His subsequent career, both in academia and industry, has focused on the application of AI, data science and search techniques to design and manufacturing. In 2006 he founded ShapeSpace Ltd, a spinout from the University of Edinburgh, to commercialise 3D search-by-shape technology, initially developed as a 3D search engine for components and subsequently enhanced to allow analysis of assemblies. This technology has been deployed at a number of large manufacturers in automotive, aerospace and industrial equipment industries where the ability to understake analytics on component portfolios and large numbers of bills of materials has uncovered significant cost savings within the supply chain. Between 2016 and 2019 he was the Royal Academy of Engineering (RAEng) Visiting Professor in Design for Product Profitability at the University of Edinburgh.

**Nan Yu** Senior lecturer in Digital Manufacture at the Institute for Materials and Processes, the University of Edinburgh (UoE). Before joining UoE, he was trained at Cranfield University (Ph.D.: plasma figuring of large optics), University College Dublin (UCD, as a postdoc in precision manufacturing of medical devices), and the European Organisation of Nuclear Research (CERN, as an Associate Scientist in precision alignment and metrology). His research interests include digitally supported process, precision manufacturing and metrology, plasma technologies. He receives the prestigious Marie Curie International Fellowship (2018−2020), Irish Research Council Fellowship (2020), and Royal Academy of Engineering Industrial Fellowship (2023−2025).