1 *Article*

2 Recovering Building Information Model from 2D
3 Drawings for Mechanical, Electrical and Plumbing
4 Systems of Ageing Buildings

5 **Pan, Zaolin[a] Yu, Yantao[a]* Xiao, Fu[b, c] Zhang, Jing[b]**

6 [a] Department of Civil and Environmental Engineering, The Hong Kong University
7 of Science and Technology, Hong Kong, China

8 [b] Department of Building Environment and Energy Engineering, The Hong Kong
9 Polytechnic University, Hong Kong, China

10 [c] Research Institute for Smart Energy, The Hong Kong Polytechnic University,
11 Hong Kong, China

12 * Correspondence: ceyantao@ust.hk

13

14 **Abstract:** Ageing buildings have become a significant concern for many cities,
15 exacerbated by inadequate management and maintenance of Mechanical, Electrical,
16 and Plumbing (MEP) systems, and Building Information Modelling (BIM) enables
17 efficient MEP Operation and Maintenance (O&M) through the digital representation
18 of system information; however, many ageing buildings were constructed without
19 BIM, and manual reconstruction is costly and inefficient due to the sheer number of
20 such structures. Although some studies have proposed methods for automatically
21 recovering BIM from 2D drawings, few are suitable for MEP systems due to the
22 multiscale and irregular shapes of MEP components. To fill this gap, an automatic
23 approach is proposed for recovering MEP BIM from 2D drawings with three modules:
24 1) semantic extraction by combining image cropping with Cascade Mask R-CNN to
25 detect and segment multiscale, irregular MEP components; 2) geometric extraction by
26 semantic-assisted image processing to extract contours and skeletons of irregular parts;
27 and 3) Industry Foundation Class (IFC)-based BIM reconstruction via the open-source
28 pythonOCC and IfcOpenShell. The performance was tested on two MEP systems with
29 335 and 282 multiscale and irregular elements, and the results show that the method
30 recovered BIMs for the two MEP systems in 2.85 seconds and 0.79 seconds, with
31 semantic extraction accuracy exceeding 0.9 and geometric error below 5%. This paper
32 contributes to the existing body of knowledge by providing a semantic and geometric-
33 based approach for recovering multiscale and irregular components from 2D drawings.
34 Future studies could further improve the approach by integrating elevation drawings,
35 reconstructing abstract symbols, and aligning text-geometry.

36

37 **Keywords:** Building Information Model (BIM); Industry Foundation Classes (IFC);

Mechanical, Electrical, and Plumbing (MEP) Systems; BIM Reconstruction; Operation and Maintenance (O&M)

1. Introduction

Deteriorating energy performance of ageing buildings challenges energy conservation and emission reduction [1]. Evidence suggests that ageing buildings over 50 years old in Europe account for 35% of existing buildings and are responsible for 40% of energy consumption and 36% of CO2 emissions in the EU [2]. Effective Operation and Maintenance (O&M) of Mechanical, Electrical, and Plumbing (MEP) systems may reduce the energy consumption and greenhouse gas emissions of an ageing building. It is estimated that up to 16% of MEP systems' energy consumption during building operations can be saved through proper management [3].

Traditional MEP system O&M methods typically retrieve information through manual review of as-built drawings and documents [4], which are complex and error-prone [5]. Building Information Model (BIM[1]) contains rich information on geometry and semantics and can improve and simplify the information retrieval process by providing practical data support for the O&M of the MEP systems [6]. The data and knowledge extracted and discovered from BIM, such as the MEP-related entities and relationships [7], the MEP logic chain [8], and the grouping and labelling of MEP equipment [9], can improve the quality and efficiency of O&M. However, most MEP systems in ageing buildings do not have BIM [10] due to the time they were constructed. Besides, manual BIM reconstruction is tedious and time-consuming. Therefore, in order to promote more efficient O&M of the many ageing buildings in the world's infrastructure, an efficient automated method for producing BIM of existing buildings is needed.

BIM reconstruction has been extensively studied. Existing studies have primarily used two types of data sources, i.e., point clouds and 2D drawings.

Point clouds, which can be acquired by depth cameras, Light Detection and Ranging (LiDAR), stereo cameras, radar, or Unmanned Aerial Vehicle (UAV) tilt photography, can capture accurate as-is building conditions. A typical process of point cloud-based BIM reconstruction methods is scan-to-BIM, which focuses on object recognition and geometric modelling [11] using various approaches, such as scan-vs-BIM [12,13] or deep learning methods [14,15]. Compared to manual modelling, scan-to-BIM significantly accelerates the BIM reconstruction process [11]. Point clouds contain rich geometric information but can only provide limited non-geometric information. Detailed parameters of MEP components, such as the boiler's heating capacity and the pump's volume flow rate, which are necessary for O&M [16], are unavailable in the point cloud. In addition, the MEP system components may be hidden in the walls, floors, or suspended ceilings, making it difficult to reconstruct these obscured parts [17]. Moreover, the acquisition and processing of point clouds require a considerable workload and cost [18].

---

[1] To reduce ambiguity, BIM in this study represents the BIM model.

78　　　Compared to point clouds, 2D drawings provide another cost-effective and
79　reliable option for BIM reconstruction. As one of the deliverables after project
80　completion, 2D drawings are easy to access and enable intact BIM reconstruction [17].
81　2D drawings contain holistic project information, including detailed components'
82　geometry and semantics, which contribute to information-rich 3D models [17] and the
83　O&M of MEP systems. In addition, as 2D drawings are not susceptible to information
84　loss due to occlusion, the reconstructed 3D models can be used as a prior model to
85　assist in the point cloud-based methods (e.g., by reducing the search space for object
86　recognition [19]). Although the geometry reconstructed from 2D drawings may not
87　match the actual building condition perfectly due to assembly deviations [14] and
88　drawing errors, this study continues to consider 2D drawings a realiable data source
89　by assuming that 1) assembly deviations are acceptable as they must meet the
90　requirements of specifications and 2) drawing errors can be reduced through drawing
91　pre-processing [20]. Consequently, this study focuses on reconstructing BIM from 2D
92　drawings for MEP systems.

93　　　Substantial research [17,21,22] has been conducted on 2D drawing-based BIM
94　reconstruction. Most of these studies have concentrated on architectural and structural
95　drawings, where most components are in regular shapes. MEP drawings, on the
96　contrary, contain multiscale, irregular components. Multiple scales mean the
97　dimensions of components vary significantly. Irregularity implies that components'
98　shapes may change along the X, Y, or Z axes and cannot be accurately annotated.
99　Most studies separate geometric extraction from semantic extraction and are unable to
100　simultaneously extract accurate semantics and geometry of multiscale, irregular MEP
101　components. In addition, research on MEP drawings is scarce and unable to
102　reconstruct the irregular shape of MEP components [23]. There is still room for
103　improvement in information extraction and component modelling of irregular
104　components in MEP systems.

105　　　Considering these limitations, this study proposes a 2D drawing-based MEP
106　model reconstruction method for irregular MEP components. Unlike existing studies,
107　this study 1) allows accurate semantic and geometric extraction of multiscale,
108　irregular MEP components through instance segmentation and semantic-assisted
109　image processing; and 2) produces efficient BIM reconstruction via an open-source
110　solution. The study will provide the O&M industry with an accurate and efficient tool
111　for reconstructing the MEP model, which will help enhance the efficiency and quality
112　of the O&M of MEP systems in ageing buildings, leading to building energy
113　conservation and emission reduction.

2. Literature review

115　　　This section reviews 2D drawing-based BIM reconstruction methods from data
116　sources, information extraction, and modelling methods. Based on the review, the
117　research gaps are identified, and the most appropriate data sources and methods for

118 reconstructing the MEP model are selected.

2.1 Data sources for BIM reconstruction

120 Two primary data sources for 2D drawing-based BIM reconstruction are CAD
121 and raster drawings [17].

122 Before the BIM era, the Architecture, Engineering, Construction, and Facilities
123 Management (AEC/FM) industries had widely used CAD software for design,
124 construction, and O&M [20]. Hence, CAD drawings that store copious building
125 information can be found for many ageing buildings [10]. By stratifying the drawings,
126 CAD drawings provide accurate geometric information and enable delicate BIM
127 reconstruction [24]. However, due to diverse naming, drawing, and annotation styles,
128 information extraction from MEP CAD drawings can be challenging. It takes
129 tremendous time and effort to pre-process the drawings [17,20], design the
130 information extraction rules [20], or manually annotate the geometrical primitives
131 (e.g., vertices and lines) [25]. Besides, an as-designed model library would be needed
132 to support 3D modelling, as the modelling parameters of some elements (e.g., valves
133 and pumps) can be insufficient in the drawings.

134 Raster drawings are electronically scanned or photocopied paper drawings that
135 are easily accessible. Compared to CAD drawings, semantic extraction from raster
136 drawings is more straightforward. Well-established and promising raster image-based
137 deep learning methods, such as object detection and instance segmentation, can be
138 used for symbol detection. Annotating raster drawings is relatively easier than
139 annotating CAD drawings because symbols can be annotated holistically. On the other
140 hand, raster drawings have less accurate geometric information and suffer from severe
141 symbol obscuration and line segment overlap, as all geometrical primitives are
142 represented by discrete pixels and are placed in the same layer. Nonetheless, these
143 gaps can be closed with image processing methods [26] that allow for accurate
144 geometric extraction.

145 Motivated by the above observations, this study selects raster images derived
146 from layered CAD drawings as the data source for the following reasons. First, raster
147 images can leverage well-established raster image-deep learning methods to simplify
148 semantic extraction, reducing manual intervention and annotation costs. Second,
149 raster images exported from different layers in CAD drawings help alleviate symbol
150 occlusion and line segment overlap. Third, although the geometric information in
151 raster drawings is less accurate than in CAD drawings, this gap can be narrowed by
152 using proper image processing methods.

2.2 Information extraction from 2D drawings for BIM reconstruction

154 Geometric and semantic information about components is essential for BIM
155 reconstruction. Specifically, geometric information refers to the elements' locations
156 and shapes (or dimensions), such as the coordinates of element contours; semantic
157 information describes element characteristics, such as element types [21]. This section

158 introduces the primary method for extracting geometric and semantic information,
159 shows where current research is lacking, and then chooses and illustrates the most
160 suitable methods for extracting information from 2D MEP raster drawings.

161 Regarding geometric extraction, current studies have used various approaches to
162 recognise elements' locations, contours, and dimensions. These methods include
163 image processing, pattern recognition, and annotation analytics. Image processing
164 methods, such as the Hough circle transform [22], connected domain labelling [27],
165 and contour detection [28], can detect the accurate pixel region and contours of
166 geometrical primitives. Pattern recognition methods generally leverage pre-defined
167 rules or templates to detect the locations and shapes of specific symbols [21]. These
168 methods are effective for parts with distinctive patterns. However, they are sensitive
169 to parts with rotations and scale variations [29]. Annotation analytics methods [22]
170 obtain components' dimensions by analysing dimensional annotations. Although these
171 methods perform well for regular components, they cannot provide irregular
172 components' exact locations and shapes.

173 As for semantic extraction, most studies focus on element-type recognition. The
174 methods used can be divided into hand-designed methods and deep learning methods.
175 The hand-designed methods [30] work better for specific divisible components, e.g.,
176 diffusers and Variable Air Volume (VAV), and produce high-quality results. However,
177 designing features or rules requires considerable time and expertise, and their
178 generalisation and robustness are poor [31]. These problems can be solved by deep
179 learning methods such as object detection, semantic segmentation, and instance
180 segmentation, which use neural networks to extract features [17]. Object detection
181 methods [17,31] are more effective for divisible instances while being less effective
182 for amorphous components (e.g., ducts) [25]. Compared to object detection, semantic
183 segmentation [32] has better recognition performance for amorphous components,
184 capable of capturing the outer contour of elements. Their shortcoming is their
185 inability to distinguish between different instances. Instance segmentation [33] can
186 detect elements' bounding boxes and outer contours. They have better support for
187 divisible and amorphous components. The problem is that the contours detected do
188 not exactly match the actual shape, as they do not sufficiently consider the geometric
189 information, resulting in rough boundaries [33].

190 Depending on the information extraction process, existing studies can be divided
191 into two paradigms. The first detects components' geometrical primitives using image
192 processing methods, then leverages hand-designed methods to recognise the
193 semantics of these geometrical primitives [21]. The second uses deep learning
194 methods to recognise more robust semantics and infer geometry from dimensional
195 annotations [17]. However, these studies, which used deep learning and image
196 processing methods separately, were unable to obtain accurate semantics and
197 geometry of multiscale, irregular MEP components simultaneously. While image
198 processing methods provide accurate geometry of multiscale, irregular MEP
199 components, hand-designed methods do not enable robust semantic extraction.

200 Although deep learning methods can robustly recognise the semantics of amorphous
201 MEP components, dimensional annotations can provide limited geometric information.
202   Theoretically, integrating instance segmentation with image processing methods
203 supports accurate semantic and geometric extractions, because 1) the instance
204 segmentation methods can provide robust semantics and rough object contours, which
205 help reduce the interference from other irrelevant components in geometric extraction,
206 and 2) image processing methods can compensate for the deficiency of instance
207 segmentation methods in geometric extraction. As a result, for information extraction,
208 this study employs instance segmentation and image processing methods.

2.3 Data schema for BIM reconstruction

210   A BIM data schema is essential for organising and storing the information
211 extracted from drawings. Revit and Industry Foundation Classes (IFC) files are the
212 current mainstream BIM file formats, capable of storing rich geometric and semantic
213 information and allowing information management during the entire lifecycle of an
214 AEC/FM project [34]. Although many engineers use Revit, its files are not open
215 source. An additional data schema conversion process is required to ensure data
216 exchange between different BIM software. Compared to Revit files, IFC is an open
217 and vendor-neutral BIM data schema that supports efficient data exchange and
218 collaboration among multi-professional participants. Considering the cross-party
219 involvement in the O&M of ageing buildings, this study chooses the IFC schema as
220 the carrier of extracted BIM data.
221   Most previous IFC-based BIM reconstruction studies focused on regular elements
222 in architectural and structural drawings, such as beams and columns. They produced
223 BIM reconstruction by combining element semantics with dimensional annotations
224 [17,22] or extruding element contours along the Z-axis [21,24]. However, these
225 modelling paradigms are not fully applicable to MEP drawings, because dimensional
226 annotations cannot provide accurate geometric information for irregular components,
227 and component shapes may change along the Z-axis. In addition, existing MEP
228 drawing-based studies modelled MEP components except for ducts as rectangular
229 bodies, ignoring the actual components' contours [20,23]. However, different contours
230 of components may represent different physical functions. There is room for
231 improvement when it comes to modelling irregular MEP components.

2.4 Research gaps

233   It is observed that most 2D drawing-based MEP model reconstruction methods 1)
234 separate the image processing-based geometric extraction from the deep learning-
235 based semantic extraction and cannot simultaneously extract the accurate semantics
236 and geometry of multiscale, irregular MEP components; and 2) are unable to model
237 the irregular shapes of MEP components. As a result, the accurate and information-
238 rich 3D MEP model cannot be reconstructed, affecting the BIM application in the
239 O&M of MEP systems.

3. Research objectives and questions

This research aims to reconstruct building information models from 2D drawings for MEP systems. The research objectives are (1) to automatically extract the semantic and geometric information of the MEP components from 2D drawings and (2) to create an IFC-based MEP model with accurate shapes preserved. To achieve these objectives, the subsequent questions will be answered.

First, how to extract accurate semantic information about the irregular and multiscale MEP components in oversized drawings? Although the instance segmentation model supports divisible and amorphous component recognition, accurate semantic recognition for multiscale components from oversized images using the instance segmentation model would be challenging. For example, if a large-size raster image (e.g., 5480 x 2880) was used directly for training, it would need to be rescaled to a smaller size (e.g., 600 x 600) for input into the model. This process will cause most of the detailed information to be lost, leading to poor detection performance for small-scale objects. If, on the other hand, small image patches were used to train models, large-scale objects would be harder to detect because an image patch only shows a small part of a large-scale object.

Second, how to extract accurate geometric information about the irregular and amorphous MEP components? While the instance segmentation model can provide component contours, the irregular component contours detected by the instance segmentation model can be inaccurate. Fig. 1, for example, shows that some vertices are missing, straight lines are detected as curves, and the detected mask does not fully cover the element's shape. These issues will make distinguishing the geometric primitives of different components extremely difficult, further affecting the accuracy of the 3D shapes created.
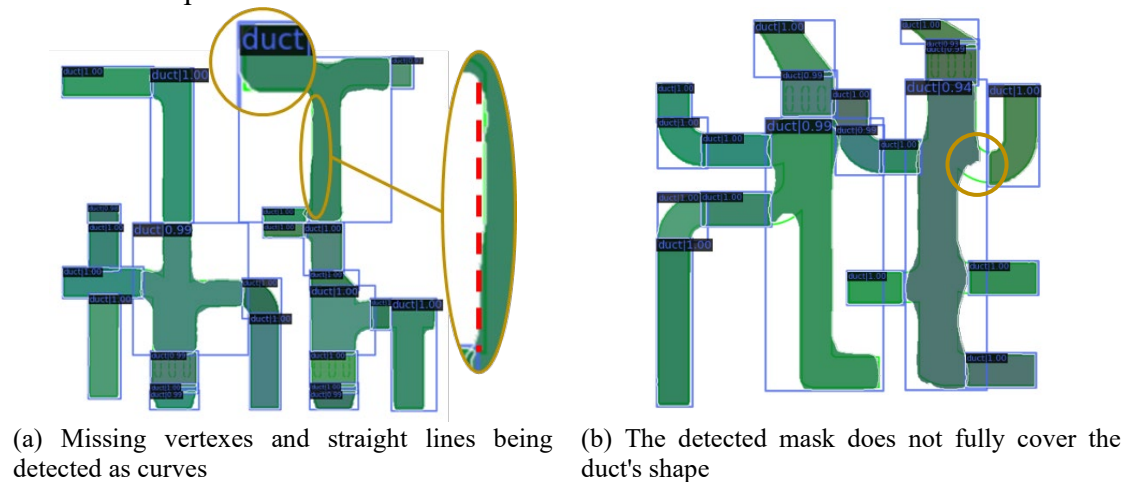

(a) Missing vertexes and straight lines being detected as curves

(b) The detected mask does not fully cover the duct's shape

Fig. 1. A result of instance segmentation.

Third, how to model the IFC-based BIM for irregular MEP components with their accurate shapes preserved? Currently, IFC modelling of irregular MEP components can be challenging as there is a lack of tools to efficiently model components with irregular shapes. If irregular components are modelled based on the IFC schema, the

component's vertex coordinates need to be specified manually. For components whose shape changes in the Z-direction, such as a hose composed of triangular faces, it is non-trivial to obtain their vertex coordinates, which will hinder the modelling efficiency of irregular components.

To answer these questions, a 2D drawing-based BIM reconstruction framework for MEP systems is proposed. This framework is comprised of three modules: semantic information extraction, geometric information extraction, and BIM automatic modelling. In the following sections, the technical particulars of the three modules will be presented.

4. Methodology

The proposed three-module framework is shown in Fig. 2. In particular, the locations, shapes, and types of MEP components constitute the geometric and semantic information in this study. A strategy for image cropping is implemented in the semantic information extraction module to facilitate multiscale object recognition in oversized drawings. This strategy can improve the identification accuracy of multiscale objects by using cropped images for training and mapping the detection results onto the original image for testing. In the geometric information extraction module, two novel semantic-assisted image processing methods for fine-grained geometric extraction of irregular components are developed. These methods, which match instance segmentation results with geometrical primitives detected by image processing techniques, can accurately distinguish the geometrical primitives of irregular components. In the BIM automatic modelling module, pythonOCC[2] and IfcOpenShell[3] are introduced to automate IFC-based BIM reconstruction for irregular MEP components. IfcOpenShell is an open-source IFC library and geometry engine, whereas pythonOCC is an open-source full-size 3D geometry library. The combination of pythonOCC and IfcOpenShell can model irregular element shapes with only a few parameters, greatly enhancing irregular components' modelling efficiency.
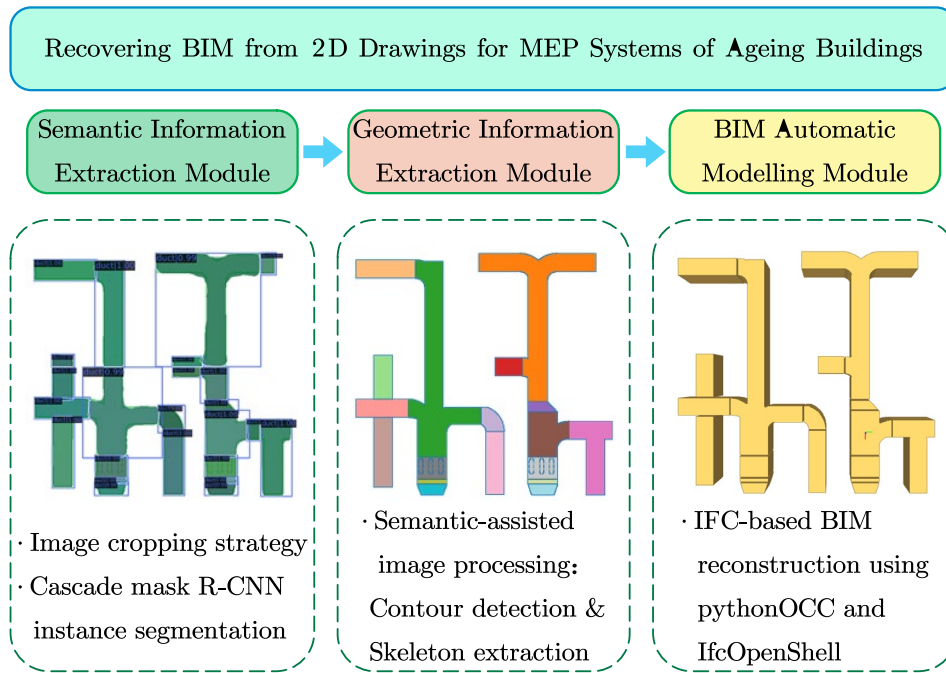
---

[2] https://github.com/tpaviot/pythonocc-core
[3] http://ifcopenshell.org

Fig. 2. The roadmap of the proposed framework.

4.1 Semantic information extraction

The semantic information extraction module aims to extract multiscale irregular MEP components' semantic information from large-size images. The following section will introduce the strategy and instance segmentation model used.

4.1.1 Image cropping strategy for multiscale component recognition

The image cropping strategy proposed by Eten [35] is adopted to improve multiscale object instance segmentation of large-size raster drawings. This strategy that segments large-size images into small-size images using sliding windows can reduce the image size and object scale differences, as shown in Fig. 3. In both the training and testing phases, the sliding window slides from left to right and from top to bottom on the image with a given overlap ratio (20% by default, as suggested by Wang et al. [36]) to segment the original image as well as the annotations. In the test phase, after non-maximum suppression processing, small-size image detection results are mapped onto the original image. This strategy preserves image details, increases training samples, and reduces annotation workload, which helps enhance model performance on large-size images and multiscale objects.
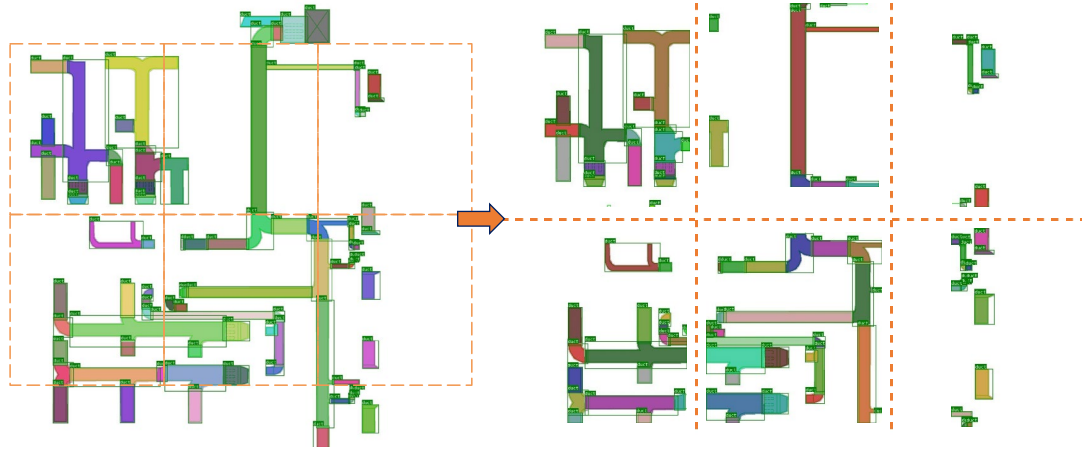
316   **Fig. 3.** Schematic diagram of using sliding windows to split images and annotations.

317   4.1.2 Cascade Mask R-CNN instance segmentation

318   The Cascade Mask R-CNN [37] model is a two-stage instance segmentation
319   model with high semantic recognition accuracy suitable for current tasks with multi-
320   category and high positioning requirements. The main feature of the Cascade Mask R-
321   CNN model is that it introduces a cascade detector. This detector allows object
322   detection and semantic segmentation at different Intersection-over-Union (IoU)
323   thresholds in different stages, significantly enhancing detection efficiency and
324   accuracy.

325   The Cascade Mask R-CNN model consists of four modules: a backbone network,
326   a Feature Pyramid Network (FPN), a Region Proposal Network (RPN), and a cascade
327   detector. The backbone network, pre-trained on a generic image dataset, can extract
328   the essential features. FPN fuses features at different scales, improving the multiscale
329   object detection capability of the network. RPN discriminates between foreground
330   and background and generates region proposals. The cascade detector contains multi-
331   stages and predicts object semantics, bounding boxes, and object masks for each stage.
332   The detector in each stage has a different IoU threshold, which can filter different
333   negative samples. By integrating these modules, the Cascade Mask R-CNN model
334   enables reliable instance segmentation for tasks with many categories and high
335   localisation requirements. Combined with the image cropping strategy, the Cascade
336   Mask R-CNN model can robustly detect and segment irregular MEP components in
337   large-size images, providing bounding boxes, contours, and semantic information for
338   geometric extraction and IFC modelling.

339   4.2 Semantic-assisted geometric information extraction

340   In order to obtain fine-grained geometry and distinguish the geometric primitives
341   of irregular elements, this module benefits from semantic information and develops
342   two semantic-assisted image processing methods for contour detection and skeleton
343   extraction, respectively. The detailed processes and the coordinate system conversion
344   method will be described in the following sections.

345 4.2.1 Semantic-assisted contour detection

346   Obtaining accurate contours is critical for modelling irregular components,
347 especially those whose shape remains constant along the Z-direction, such as ducts
348 and VAV, as they can be modelled by extruding the contours vertically. Based on the
349 results of the instance segmentation, the geometric information extraction module
350 uses a semantic-assisted contour detection method to obtain the exact shape of
351 irregular MEP components. The specific process is presented as follows. First, the
352 module merges all the bounding boxes and masks of the same class that the instance
353 segmentation model detects, as shown in the blue boxes and green and purple masks
354 in Fig. 4(a). Bounding box merging means creating a maximum bounding box by
355 calculating the minimum and maximum coordinates from the boxes with scores above
356 a given threshold (e.g., 0.6). Then, the original image can be segmented by the
357 merged bounding box. This step helps reduce the interference of other extraneous
358 components with contour detection. Afterwards, the module marks the connected
359 domains on the segmented images using a connected domain labelling algorithm. A
360 connected domain is a collection of adjacent pixels with the same intensity. For
361 example, different coloured regions represent different connected domains, as shown
362 in Fig. 4(b). This step aids in the separation of geometric primitives from different
363 components. Finally, for each connected domain, the module counts the proportion of
364 pixels within the connected domain that overlap with the position of the merged mask.
365 If this proportion exceeds the specified threshold, the connected domain is retained,
366 and contour detection is performed; otherwise, the connected domain is excluded.
367 This step further eliminates other irrelevant components and distinguishes the
368 geometric primitives of different components. Eventually, the accurate contour and
369 semantic information of components can be obtained by matching their connected
370 domains and masks.

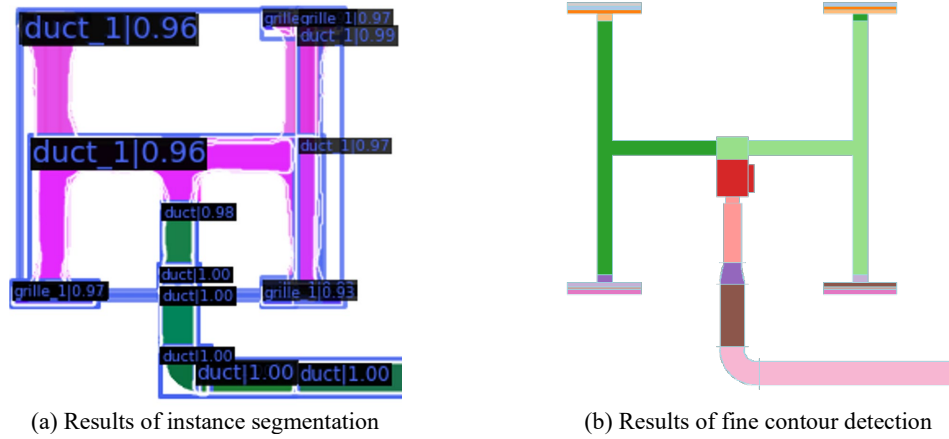(a) Results of instance segmentation   (b) Results of fine contour detection
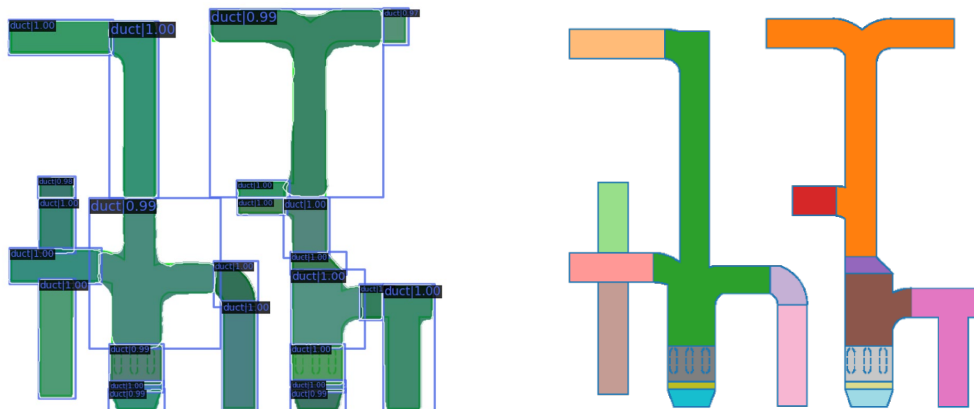
Fig. 4. Contour detection of ducts.

4.2.2 Semantic-assisted skeleton extraction

Another challenge is to extract the geometric information of components whose shapes change in the Z-direction, such as pipes and hoses. Apart from the contour information, modelling these components requires accurate skeleton information. Consequently, a semantic-assisted skeleton extraction method is developed. Specifically, the centerlines of these components are first detected based on their rough contours using the skeleton extraction method, i.e., OpenCV's cv2.ximgproc.thinning function. However, the starting and ending points of the centerline detected may be inaccurate, as it achieves thinning by iteratively shrinking the pixels of the outer contour. As illustrated in Fig. 5(b), the detected centerline is the curve between two green points, while the target centerline should be the curve between orange points. To solve this problem, this study assumes that the slope of the line segment between the orange and green dots is constant, as the distance between these two dots is close to the pipes' radius. The exact starting and ending points of the centerline can then be determined by extending the line segment to the bounding box. After locating the centerlines, these components can be modelled easily with their diameters.
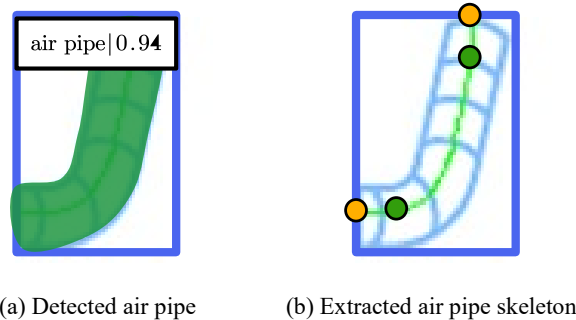

(a) Detected air pipe   (b) Extracted air pipe skeleton

Fig. 5. Air pipe skeleton extraction.

4.2.3 Coordinate system conversion

The detected pixel coordinates of components' contours and centerlines in previous sections cannot be used directly for modelling, as they do not correspond to the dimensions and locations indicated on the drawings. Therefore, it is necessary to convert the Pixel Coordinate System (PCS) to the Drawing Coordinate System (DCS).

395 This conversion requires the determination of the DCS origin location and the pixel
396 scale. For the former, this study employs the Hough circle transform and Optical
397 Character Recognition (OCR) techniques to detect the locations and semantics of axis
398 labels. The Hough circle transform is a feature extraction method to detect circles in
399 images. OCR is a technique for detecting the text location and transforming the
400 detected text into computer-encoded text. Afterwards, the DCS origin position (u_0, v_0)
401 can be inferred from the pixel coordinates of the smallest axis labels in the X and Y
402 directions, as shown in Fig. 6. As for the pixel scale calculation, the inter-axis
403 annotations L_x and L_y are first identified using OCR on the area enclosed by the axis
404 network, as shown in the light blue area in Fig. 6. The pixel scale can then be
405 calculated using the following equation after measuring the pixel distances P_x and P_y
406 between the axis labels:

407 $$S_x = \frac{\sum_{i=1}^{n} L_{x_i}}{\sum_{i=1}^{n} P_{x_i}}, S_y = \frac{\sum_{i=1}^{m} L_{y_i}}{\sum_{i=1}^{m} P_{y_i}}, \qquad (1)$$

408 where *n* and *m* are the numbers of inter-axis labels on the X- and Y-directions,
409 respectively. Equation (1) calculates the ratio of each pixel to the annotation length by
410 dividing the grid annotation length by the pixel number between grids. With the DCS
411 origin location and the pixel scale being calculated, the PCS can be converted to DCS
412 using the equation below:

413 $$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} S_x & 0 & -u_0 S_x \\ 0 & -S_y & v_0 S_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \qquad (2)$$

414 Equation (2) first converts the PCS origin $\boldsymbol{O}_{\text{PCS}}$ to the pixel coordinates (u_0, v_0) where
415 the DCS origin $\boldsymbol{O}_{\text{DCS}}$ is located, and then scales the pixel coordinates to drawing
416 coordinates using pixel scales S_x and S_y in the X- and Y-directions, respectively.
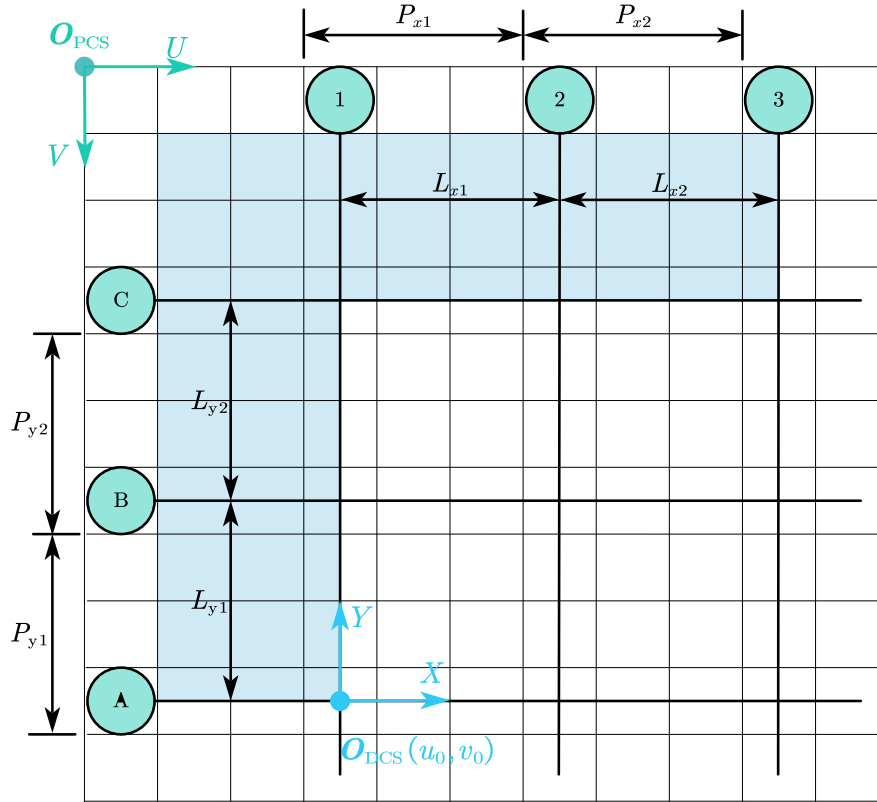417

Fig. 6. Schematic diagram of pixel coordinate system and drawing coordinate system.

4.3 BIM automatic modelling

After extracting the semantic and geometric information of components and converting the PCS to DCS, the BIM automatic modelling module can create an IFC-based BIM file using pythonOCC and IfcOpenShell. Specifically, two main steps are (a) matching the semantics of the components with IFC entities and (b) generating a geometric representation of the components. In the first step, this module matches components with IFC entities using the mapping rule shown in Table 1. In the second step, the module generates the geometric representation of components using different entities to represent the components' geometry, as illustrated in Fig. 7. For components whose shape remains constant along the Z-direction, the module uses the extrusion method. This method creates an IfcExtrudedAreaSolid entity to represent the components' geometry using detected contours and pre-defined extrusion lengths and directions. For components whose shape changes in the Z-direction, such as air terminals and hoses, this module first uses the pythonOCC engine to generate the 3D shape of these components. For example, an air inlet is considered a combination of a wedge and a box; therefore, the air inlet can be easily modelled using the BRepPrimAPI_MakeBox and BRepPrimAPI_MakeWedge functions of the pythonOCC engine. These 3D shapes in boundary representation format are then converted by IfcOpenShell to a geometric representation using the IfcFaceBasedSurfaceModel entity. It is worth noting that since the plan drawing only provides the X- and Y-axis coordinates, the Z-axis coordinates must be set manually

441   to avoid collisions.

442

Table 1. Component types and IFC entities mapping table.

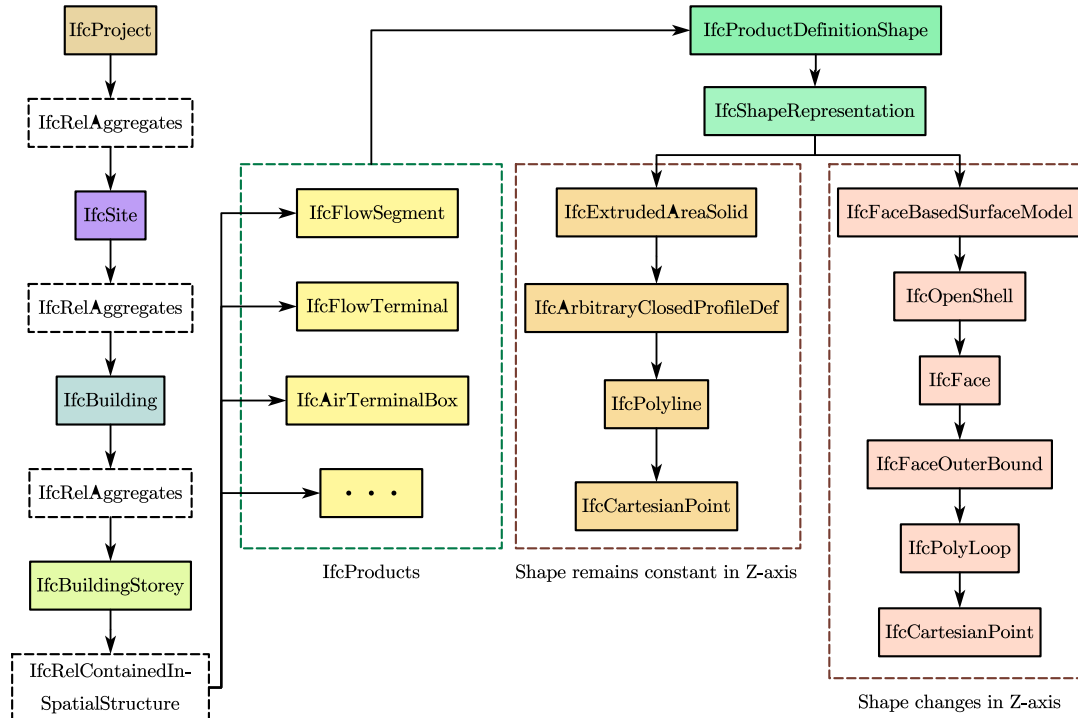| Component type	IFC entity
Duct and hose	IfcFlowSegment
Air terminal and grille	IfcFlowTerminal
FCU, AHU, muffler, and VAV	IfcAirTerminalBox

443

444

445   **Fig. 7.** Schematic diagram of the IFC file structure.

5. Experiment

447   To verify the feasibility and reliability of the proposed method, 2D MEP
448 drawings from existing buildings were used for the experiment. The detailed
449 experimental setup and results are presented in the following sections. The code is
450 available on GitHub (https://github.com/CrossStyle/MEP2BIM).

5.1 Experimental setup

452   This section elaborates on the experimental data, environment, parameters, and
453 evaluation metrics.

5.1.1 Experimental data

455   A dataset of 2D MEP drawings is created, including 335 images. These images
456 are cropped from nine raster images exported from different layers in CAD drawings
457 of two existing buildings. The training and test sets are divided from these cropped
458 images with a ratio of 3:1. Specifically, the layer information in CAD drawings is
459 utilised to mitigate the overlap of symbols and line segments. Fig. 8 shows that the
460 number of mutually occluded components in the drawings is significantly reduced
461 after separating different layers. The tools used in this conversion are AutoCAD and

Ezdxf (https://github.com/mozman/ezdxf), which convert DWG drawings to DXF files and then to PNG images. The line width in CAD drawings is set to the minimum line width, i.e., 0.00mm in AutoCAD, to reduce the error caused by the line width. Afterwards, these images are cropped from left to right and from top to bottom with an overlap ratio of 20% by default to provide diverse training and test data. The size of the sliding window is 600 by 600. It is worth noting that the overlap ratio is adjusted accordingly due to class imbalances and varying graphical densities. For example, the ratio is set higher (such as 50%) for classes with small sample sizes or drawings with more dense graphical primitives.

The statistics of the number and area of each component type are shown in Table 2. These components can be grouped into three categories depending on their size. Specifically, ducts are classified as large-size components, as they account for only 27.0% of the quantity but occupy 78.2% of the area. In contrast, Fan Coil Units (FCUs), VAV, grilles, and air terminals are small-size components, as their area shares are much lower than their quantity shares. The rest of the components, i.e., mufflers and Air Handling Units (AHUs), are medium-sized. Their area shares are slightly larger than their quantity shares.
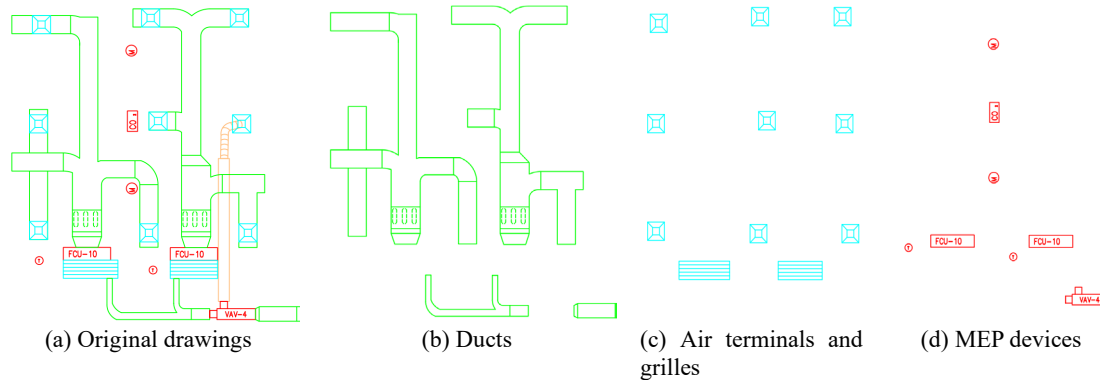

| (a) Original drawings | (b) Ducts | (c) Air terminals and grilles | (d) MEP devices |

Fig. 8. Exported drawings in layers.

Table 2. The quantity and area of each type of component.

Type	IFC entity	Scale	Quantity	Area (m^2)	Quantity (%)	Area (%)
Duct	IfcFlowSegment	Large	72	224.680	0.270	0.782
Muffler	IfcAirTerminalBox	Medium	2	2.942	0.007	0.010
AHU	IfcAirTerminalBox	Medium	2	4.065	0.007	0.014
FCU	IfcAirTerminalBox	Small	18	6.111	0.067	0.021
VAV	IfcAirTerminalBox	Small	27	11.555	0.101	0.040
Grille	IfcFlowTerminal	Small	98	25.843	0.367	0.090
Air terminal	IfcFlowTerminal	Small	48	12.000	0.180	0.042

5.1.2 Experimental environment and parameters

The experiment environment is Windows 11, 64 G of RAM, a 3070 laptop graphics card, an i7-10870H processor, Python 3.9, and PyTorch 1.8.1. The total number of training epochs is 12, and the learning rate is reduced to one-tenth of the previous one in the 8th and 11th epochs. The optimiser is stochastic gradient descent with a momentum of 0.9 and a weight decay of 0.00001.

488 5.1.3 Evaluation Metrics

489 Different evaluation metrics are adopted to evaluate the accuracy of semantic
490 extraction, geometric extraction, and component modelling.

491 For semantic extraction, this study chooses Average Precision (AP), which can
492 evaluate the localisation and segmentation accuracy of the instance segmentation
493 model at different IoU thresholds and scales. Specifically, AP_{50} and AP_{70} represent
494 the AP value when the IoU threshold is 0.5 and 0.7, respectively. The mean Average
495 Precision (mAP) is the average of the AP calculated when the IoU threshold is 0.5 to
496 0.95 with a step size of 0.05. AP_S, AP_M and AP_L represent the AP values for small
497 targets (pixel area less than 32×32), medium targets (pixel area between 32×32 and
498 96×96), and large targets (pixel area greater than 96×96), respectively. The procedure
499 for calculating AP is shown in the following equation:

$$AP = \int p(r)dr, \tag{3}$$

501 where p represents the precision, r represents the recall, and AP is the area enclosed
502 by the p-r curve and the coordinate axis.

503 The evaluation metrics for geometric extraction and component modelling are
504 perimeter and area, as they allow straightforward evaluation of the shape generated
505 from 2D drawings. The calculation procedure is shown below:

$$Er_A = \frac{A - A_{\mathrm{GT}}}{A_{\mathrm{GT}}}, A = A_{\mathrm{pixel}} \times N_{\mathrm{pixel}}, \tag{4}$$

$$Er_P = \frac{P - P_{\mathrm{GT}}}{P_{\mathrm{GT}}}, \tag{5}$$

508 where A and P denote the area and perimeter, respectively; A_{GT} and P_{GT} represent the
509 ground truth values of A and P, respectively; Er_A and Er_P denote the error of the
510 generated components' area and perimeter, respectively. A can be calculated by
511 multiplying the pixel number N_{pixel} in the detected connected domain by the DCS
512 area of each pixel A_{pixel}. P can be calculated by accumulating the distance between
513 vertices in the detected contour.

5.2 Experimental results

515 This section presents the experimental results of three tasks, i.e., semantic
516 information extraction, geometric information extraction, and IFC modelling.

517 5.2.1 Semantic information extraction results

518 The training loss, training accuracy, and test mAP are shown in Fig. 9. In the first
519 400 iterations, loss declined dramatically, and the accuracy of two stages of the
520 Cascade Mask R-CNN model (i.e., S1 acc and S2 acc) in the training phase and the
521 mAP of two tasks (i.e., object detection and instance segmentation) in the test phase
522 rose rapidly. Although the magnitude of their changes became smaller in the
523 following iterations, the loss was still going down and both mAPs were still going up,
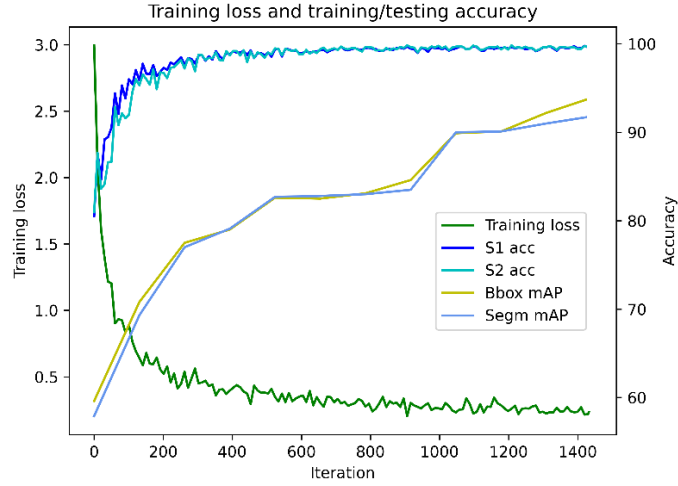524 which shows that there was no overfitting.

Fig. 9. The diagram of training loss, training accuracy, and test mAP. S1 acc and S2 acc represent the accuracy of the two stages of the Cascade Mask R-CNN model in the training phase. Bbox mAP and Segm mAP are the mAPs for the object detection and instance segmentation tasks in the test phase, respectively.

The qualitative results of the Cascade Mask R-CNN model are shown in Fig. 10. Fig. 10 shows that the model works well for detecting components of different scales and shapes, both regular and irregular.
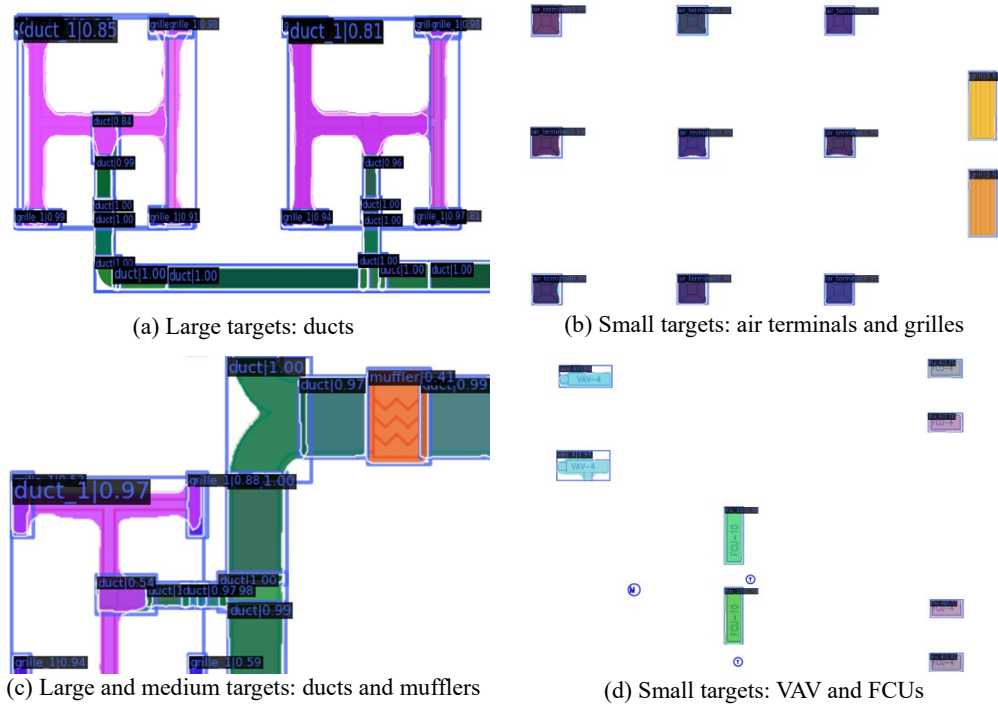

(a) Large targets: ducts

(b) Small targets: air terminals and grilles

(c) Large and medium targets: ducts and mufflers

(d) Small targets: VAV and FCUs

Fig. 10. Qualitative results of the Cascade Mask R-CNN instance segmentation model.

The quantitative results of the Cascade Mask R-CNN model are shown in Table 3 and Table 4. According to Table 3, the mAP of the model is higher than 0.9, with excellent performance in object detection and instance segmentation tasks. The model

also achieved good large- and medium-scale results, with the AP metrics greater than 0.9, followed by the small-scale components. Table 4 presents the performance of various components. The mAPs of all components' object detection and instance segmentation are greater than 0.80, indicating good overall performance. The above qualitative and quantitative results demonstrate the reliability of the Cascade Mask R-CNN model in the semantic information extraction task.

Table 3. Quantitative results of the Cascade Mask R-CNN model in different AP metrics.

Task	mAP	AP_{50}	AP_{70}	AP_S	AP_M	AP_L
Object detection	0.939	0.998	0.995	0.900	0.942	0.984
Instance segmentation	0.918	0.998	0.998	0.865	0.920	0.912

Table 4. mAP of the Cascade Mask R-CNN model in different classes.

| Task	Muffler	Duct	Duct1	Air pipe	VAV4	VAV5	FCU10	FCU8	FCU4	Air terminal	AHU	Grille	Grille1
Bbox*	0.968	0.864	0.940	1.000	1.000	0.883	0.925	0.900	0.954	0.927	1.000	0.950	0.883
Segm*	0.907	0.851	0.821	0.911	1.000	0.857	0.925	0.900	0.949	0.904	1.000	0.953	0.873

* Bbox and Segm represent the object detection and instance segmentation tasks, respectively.

5.2.2 Geometric information extraction and IFC reconstruction results

The reconstructed IFC models of two MEP systems are shown in Fig. 11. Both regular and irregular MEP system components in the 2D drawings are recognised and reconstructed. After extracting semantic and geometric information, the reconstruction durations are 2.85s for MEP systems 1 of 335 components and 0.79s for MEP systems 2 of 282 components. The reconstructed 3D models and reconstruction errors of some components are presented in Table 5. Table 5 shows that the extracted component perimeter and area errors are less than 5%. These results demonstrate the reliability, accuracy, and efficiency of the proposed geometric information and IFC reconstruction methods.
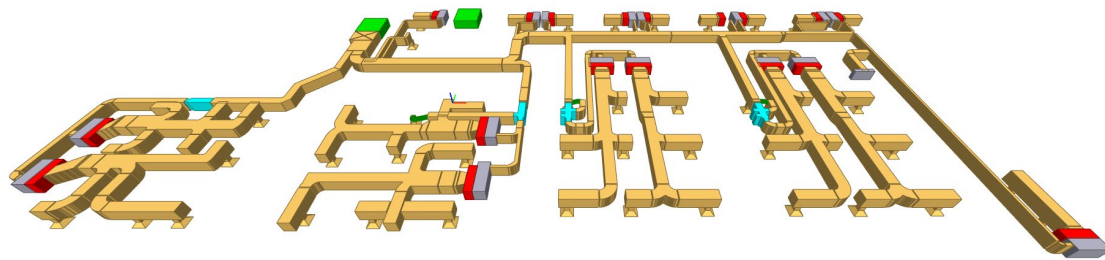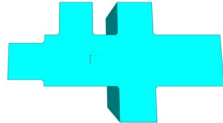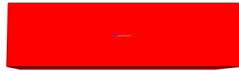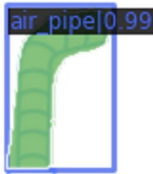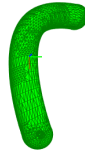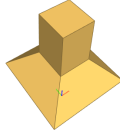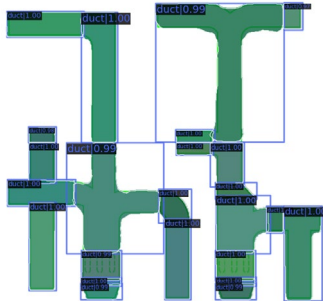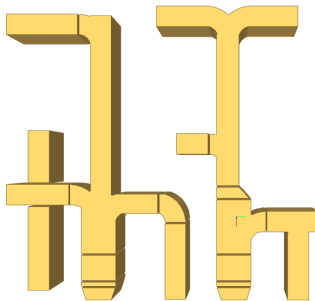

(a) Reconstructed MEP system 1
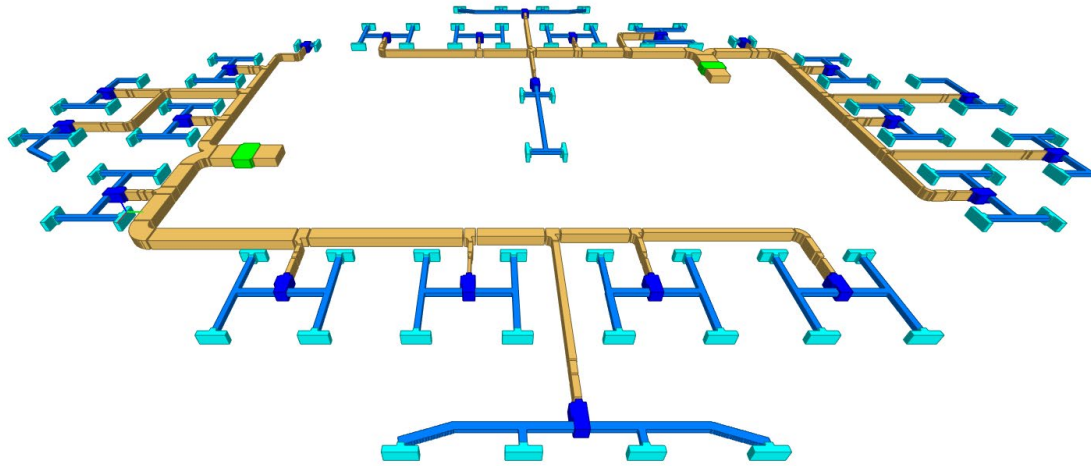
(b) Reconstructed MEP system 2

Fig. 11. Reconstructed IFC models.

Table 5. Geometric extraction and IFC reconstruction results of some components.

Type (Perimeter&Area)	Detection results	IFC reconstruction results	Errors
VAV-4 (P: 4.351m, A: 0.481m^2)			P:-0.16% A:-0.20%
FCU-10 (P: 3.300m, A: 0.455m^2)			P:-0.59% A:- 0.41%
Hose (NA)			NA
Air terminal (P: 1.200m, A: 0.090m^2)			P:±0.00% A:±0.00%
Duct (P: 63.299m, A: 16.189m^2)			P:-4.12% A:-3.03%

6. Discussion

This study provides a feasible and reliable method for reconstructing irregular

MEP models based on 2D drawings. As mentioned in the previous section, regular and irregular MEP system components in the 2D drawings can be recognised and modelled accurately and efficiently. However, there are still factors that affect or limit the information extraction and the IFC modelling process. The following section will discuss the effect of the image cropping strategy on semantic extraction and analyse the factors that cause inaccurate contours on geometric extraction and the constraints on IFC modelling.

6.1 Effect of image cropping strategy on semantic extraction

The image cropping strategy is adopted for multiscale object segmentation in the MEP drawings. To illustrate the effectiveness of this strategy, three different testing strategies are compared. Specifically, these strategies are to use (a) unsegmented images, (b) segmented non-overlapping small images, and (c) segmented partial-overlapping small images for testing. The experimental results are shown in Fig. 12 and Table 6. Fig. 12(a) shows that considerable detailed information is lost when using the unsegmented image for testing. Most of this detailed information can be retrieved using non-overlapping segmented images, as shown in Fig. 12(b). Fig. 12(c) demonstrates that all components are correctly detected. Table 6 also displays similar results. The best results were obtained using the segmented partial overlapped setting, with all AP metrics greater than 0.8. These results demonstrate that the image cropping strategy helps improve model performance on multiscale object segmentation, validating its effectiveness.
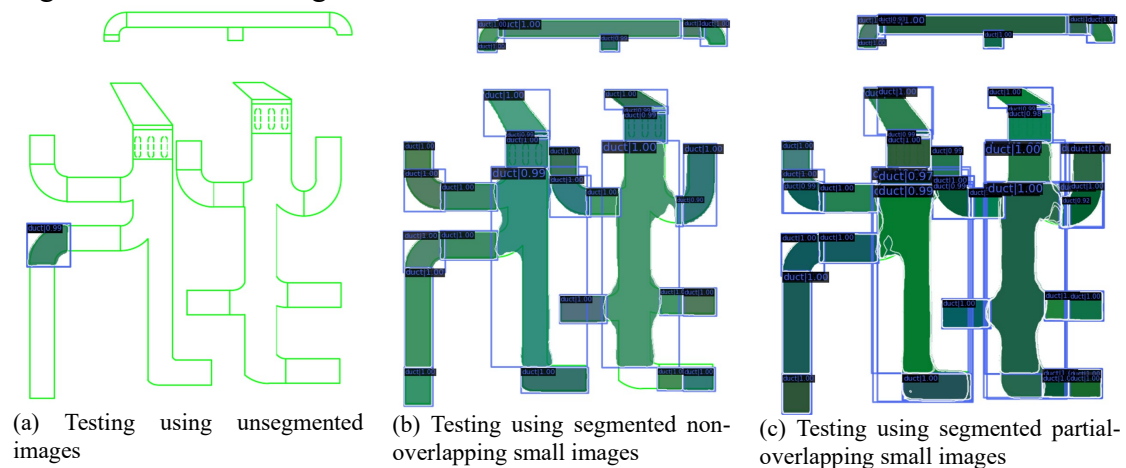

(a) Testing using unsegmented images (b) Testing using segmented non-overlapping small images (c) Testing using segmented partial-overlapping small images

Fig. 12. A comparison of three strategies with a score threshold of 0.9. The outlined green rectangles are the shapes of ducts to be detected.

Table 6. mAP comparison of three different strategies for testing.

Testing strategy	Object detection				Instance segmentation			
	mAP	AP_S	AP_M	AP_L	mAP	AP_S	AP_M	AP_L
Unsegment	0.141	0.120	0.175	0.078	0.146	0.122	0.174	0.082
Segmented non-overlap	0.886	0.759	0.882	0.952	0.857	0.740	0.849	0.894
Segmented partial-overlap	0.962	0.855	0.968	0.994	0.932	0.818	0.938	0.920

6.2 Error analysis of geometric extraction

Accurate contours contribute to accurate 3D reconstruction. However, the contours may change due to different settings of line widths and resolutions that affect the number of pixels occupied by symbols. This section will analyse and discuss the impact of line width and resolution on geometric extraction.

6.2.1 Influence of line widths on the accuracy of geometric extraction

Different line widths may result in differences in components' outer contours, thereby affecting the accuracy of the extracted geometric information. Take VAV-4 as an example, as depicted in Fig. 13. The increase in line widths leads to an increase in the number of pixels occupied by the contour, causing the contour to expand, as shown in Table 7. To evaluate the effect of this contour expansion caused by line widths, three line widths, i.e., 0.00mm, 0.25mm, and 0.50mm, are compared, where 0.00mm refers to the minimum line width in AutoCAD. The experimental results are shown in Table 7 and Fig. 14. Table 7 reveals that the errors in the area and perimeter of VAV-4 increase as the line width increases. Fig. 14 demonstrates that an increase in line width leads to inaccurate corner points in the generated IFC model. As a result, the line width of the drawings used in this study will be set to 0.00mm to reduce the error caused by the line width.
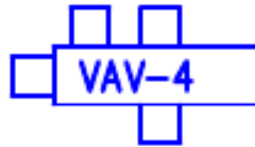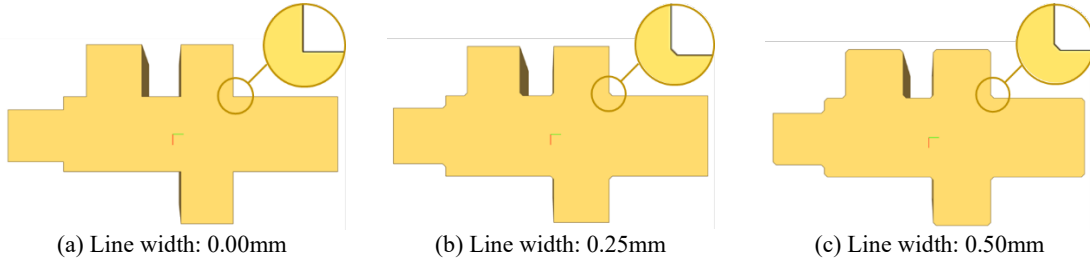

Fig. 13. A 2D drawing of VAV-4.

| (a) Line width: 0.00mm | (b) Line width: 0.25mm | (c) Line width: 0.50mm |

Fig. 14. Generated IFC models at three different line widths.

Table 7. The relative error of the area and perimeter of VAV-4 extracted at three line widths.

GT	Line width: 0.00mm	Line width: 0.25mm	Line width: 0.50mm
P (m): 4.351	4.292 (-1.33%)	4.371 (+0.48%)	4.462 （+2.57%）
A (m^2): 0.481	0.479 (-0.49%)	0.537 (+11.57%)	0.585 (+21.56%)

6.2.2 Influence of image dimensions on geometry information extraction

In addition to line widths, image dimensions also affect the number of pixels

616   occupied by the symbols. To investigate this effect, the VAV-4 depicted in Fig. 13
617   continues to be used as an example, and its original image dimensions are 102×57.
618   The VAV-4 is magnified by a factor of five and ten for comparison. The results are
619   shown in Table 8 and Fig. 15. Table 8 demonstrates that if the line width is set to
620   0.00mm, increasing the image dimensions helps reduce the error of the area and
621   perimeter. Figs. 15(a)-(c) also support this conclusion, as the IFC model generated
622   using the enlarged image has more accurate corner points. However, it is worth noting
623   that this conclusion applies only when the line width is 0.00mm. The last column of
624   Table 8 and Fig. 15(d) demonstrate that if the line width is not 0.00mm, the error will
625   be greater. Therefore, in order to improve the accuracy of detected contours in CAD
626   drawings, the image's dimensions should be appropriately scaled up after setting the
627   line width to 0.00mm. Since the line widths of paper drawings are not adjustable, the
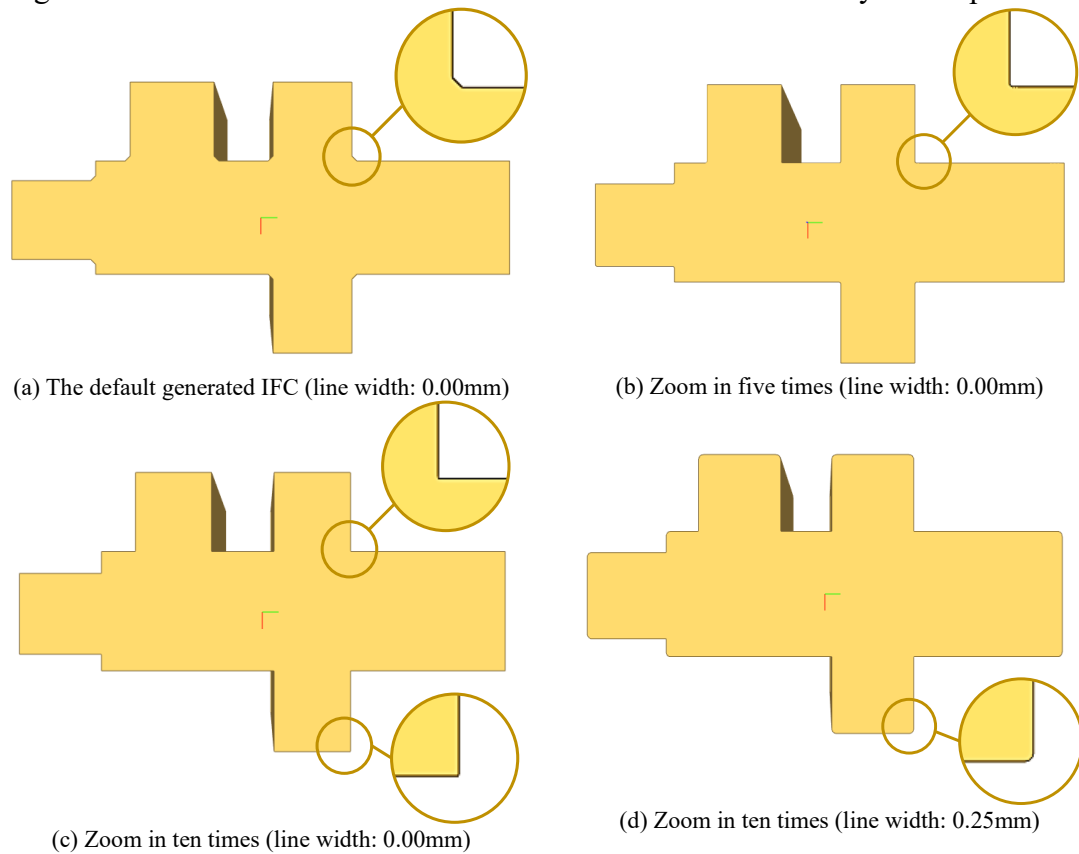628   image's dimensions should not be increased to avoid errors caused by line expansion.

| (a) The default generated IFC (line width: 0.00mm) | (b) Zoom in five times (line width: 0.00mm) |
| (c) Zoom in ten times (line width: 0.00mm) | (d) Zoom in ten times (line width: 0.25mm) |

629   **Fig. 15.** Generated IFC models at different image dimensions and line widths.

630   **Table 8.** The relative error of the area and perimeter of VAV-4 extracted at different image dimensions
631   and line widths.

GT	x1 (0.00mm)	x5 (0.00mm)	x10 (0.00mm)	x10 (0.25mm)
P (m):	4.292	4.332	4.344	4.392
4.351	(-1.33%)	(-0.42%)	（-0.16%）	（+0.94%）
A (m^2):	0.479	0.483	0.482	0.534
0.481	(-0.49%)	(+0.45%)	(+0.20%)	(+11.09%)

632   ## 6.3 Generated IFC analysis

633   Despite the promising results of the proposed method in detecting and modelling

634 irregular components, there are still constraints in converting 2D elements to IFC
635 models. This section will discuss these restrictions and the quality of the resulting IFC
636 file.

637 6.3.1 MEP elements to IFC conversion

638   Two types of components may not be converted correctly: (a) the element with a
639 similar shape but different semantics, and (b) the element whose shape in the drawing
640 does not represent its actual physical shape. Typical elements of the first type are
641 ducts and pipes. They are both represented by two parallel lines but with different
642 annotations, as shown in Fig. 16. The proposed methods cannot be applied to these
643 components because those with similar shapes will have similar visual features and,
644 therefore, the same semantics. To correctly reconstruct the shape of MEP elements,
645 the element shapes in the drawing should be the actual projection of their 3D physical
646 shapes, as the proposed method infers 3D shapes from 2D contours. However, not all
647 components meet this requirement, e.g., sensors and valves, as shown in Fig. 17.
648 These devices are generally represented by abstract symbols, which do not represent
649 their actual shapes. Furthermore, since Z-axis coordinates are not available from plan
650 drawings, they need to be set manually. However, the manually set coordinates may
651 not match the actual pose of the element that does not have constant Z-axis
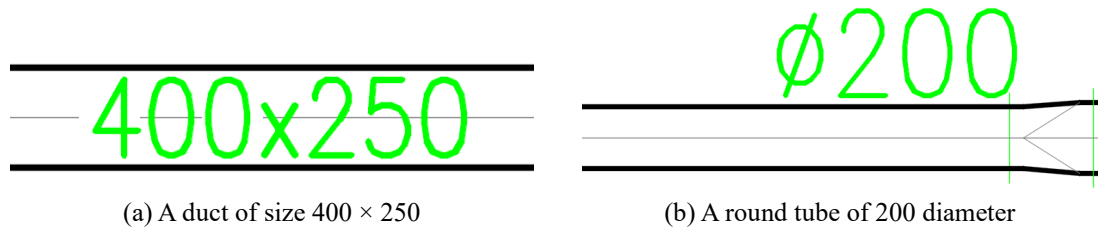652 coordinates, such as hoses. Table 9 displays the conversion of MEP components to
653 IFC models.

(a) A duct of size 400 × 250   (b) A round tube of 200 diameter

654 **Fig. 16.** Components with similar shapes but have different annotations.

655

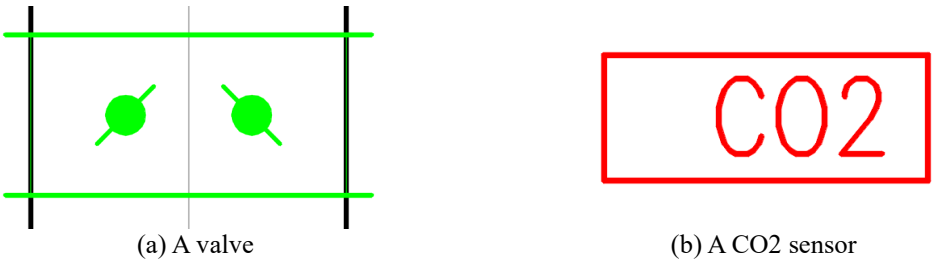(a) A valve   (b) A CO_2 sensor

656 **Fig. 17.** Two abstract symbols in the MEP drawings.

657

658 **Table 9.** The IFC conversion of MEP components.

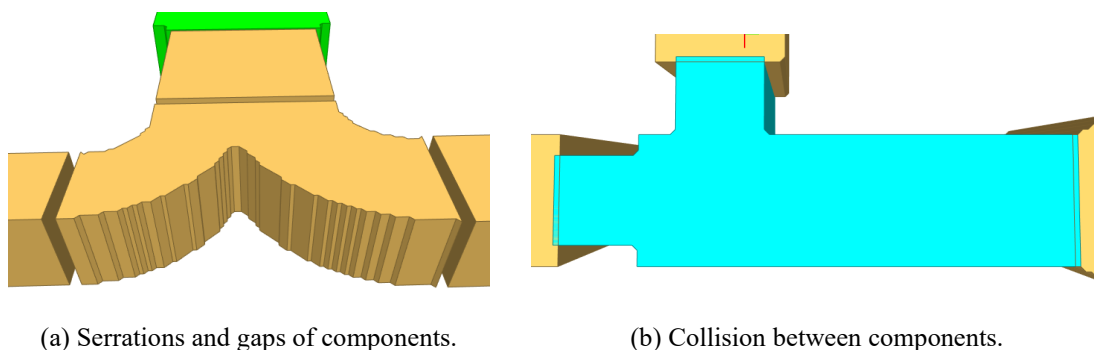| Components that can be converted with high accuracy	Components that can be converted but may have deviations	Non-convertible components
Ducts, AHUs, VAV, FCUs, grilles, air terminals, and mufflers	Hoses and tubes	Sensors, switches, valves, pumps and abstract device symbols

659

660 6.3.2 Resulting IFC file quality

661   Fig. 18 depicts a sample IFC file using the IFC4 schema. Since this study focuses
662 on semantic and geometric extraction, the object type and geometry of MEP elements
663 are included in the IFC file but not their detailed attributes in annotations. As
664 discussed in the previous section, it is possible to accurately reconstruct the
665 components whose shapes in the drawing represent their actual shapes. However,
666 incorrect cases persist. The recovered components have gaps, collisions, and serrated
667 surfaces, as shown in Fig. 19. Gaps and collisions are caused by contour lines because
668 they occupy a certain number of pixels. Errors may be introduced in these areas when
669 separating elements by connected domain. The serrated surface is the result of discrete
670 pixel coordinates because the contours created by pixel coordinates are not smooth.
671 The primary purpose of the generated IFC file is to create the logic chain of MEP
672 components based on their types, shapes, and spatial locations for fault detection and
673 diagnosis.

674
```
ISO-10303-21;↓
HEADER;↓
FILE_DESCRIPTION(('ViewDefinition [CoordinationView]'),'2;1');↓
FILE_NAME('hello_duct.ifc','2022-10-14T05:50:40',('Zaolin Pan'),('HKUST'),'IfcOpenShell','IfcOpenShell','');
FILE_SCHEMA(('IFC4'));↓
ENDSEC;↓
DATA;↓
#1=IFCPERSON($,$,'Zaolin Pan',$,$,$,$,$);↓
#2=IFCORGANIZATION($,'UST',$,$,$);↓
#3=IFCPERSONANDORGANIZATION(#1,#2,$);↓
#4=IFCAPPLICATION(#2,'0.7','IfcOpenShell','');↓
#5=IFCOWNERHISTORY(#3,#4,$,.ADDED.,$,#3,#4,1665726640);↓
#6=IFCDIRECTION((1.,0.,0.));↓
#7=IFCDIRECTION((0.,0.,1.));↓
#8=IFCCARTESIANPOINT((0.,0.,0.));↓
#9=IFCAXIS2PLACEMENT3D(#8,#7,#6);↓
#10=IFCDIRECTION((0.,1.,0.));↓
#11=IFCGEOMETRICREPRESENTATIONCONTEXT($,'Model',3,1.E-05,#9,#10);↓
#12=IFCDIMENSIONALEXPONENTS(0,0,0,0,0,0,0);↓
#13=IFCSIUNIT(*,.LENGTHUNIT.,.MILLI.,.METRE.);↓
#14=IFCSIUNIT(*,.AREAUNIT.,.MILLI.,.SQUARE_METRE.);↓
#15=IFCSIUNIT(*,.VOLUMEUNIT.,.MILLI.,.CUBIC_METRE.);↓
#16=IFCSIUNIT(*,.PLANEANGLEUNIT.,$,.RADIAN.);↓
#17=IFCMEASUREWITHUNIT(IFCPLANEANGLEMEASURE(0.017453292519943295),#16);↓
```

675   **Fig. 18.** The created IFC files.

676

(a) Serrations and gaps of components.   (b) Collision between components.

Fig. 19. Inaccurate cases in the generated IFC files.

7. **Limitations and future work**

Even though previous experiments have demonstrated that the proposed method can be used to reconstruct MEP models, the current research still has limitations.

In terms of semantic extraction, identifying the semantics of the same or similar-shaped object, e.g., ducts and tubes, is challenging. This method may therefore not be applicable to 2D drawings in which all pipes and tubes are depicted as single or parallel lines. The primary reason is that the current method only uses visual features and does not take annotations into account. Future studies may recognise the element types in the annotations to refine the detection results. Due to the lack of 2D drawings in the created dataset, the current study was unable to perform fine-grained semantic recognition, such as differentiating between ducts and connected components. Future research can collect more data to address this limitation, as a larger dataset compensates for class unbalance and improves model robustness.

Regarding geometric extraction, this study has three limitations. Firstly, due to the influence of line width and discrete pixel coordinates, the generated IFC files may contain collisions, gaps, or serrated surfaces. Future research may consider linking the pixel area in the raster drawings with the geometric primitives in the CAD drawings. Both semantics and geometry can be obtained precisely in this manner. Secondly, the current method requires manual pre-processing to correct drawing errors, such as line segments that are not properly closed or out of range, as depicted in Fig. 20. These errors result in inaccurate contours for components. Future studies may wish to develop drawing error correction methods that automatically perform topology analysis to merge the closest endpoints. Thirdly, the Z-axis coordinates are missing from the plan drawing and thus need to be set manually, resulting in incorrect 3D poses. Future studies can combine elevation drawings with other data sources (e.g., on-site images or point clouds) to extract this information.

As for IFC modelling, the proposed method cannot reconstruct those abstract symbols whose shapes in the drawings do not represent their physical shapes. Future studies may consider developing a method that predicts the modelling parameters of abstract symbols with element types as input. Besides, the resulting IFC file lacks the detailed properties of components, such as the cooling and volumetric capacity of the AHU. Thus, it cannot be used directly for simulation. Future studies can enrich this information by analysing the text in annotations, symbol descriptions, and equipment schedules with OCR methods.
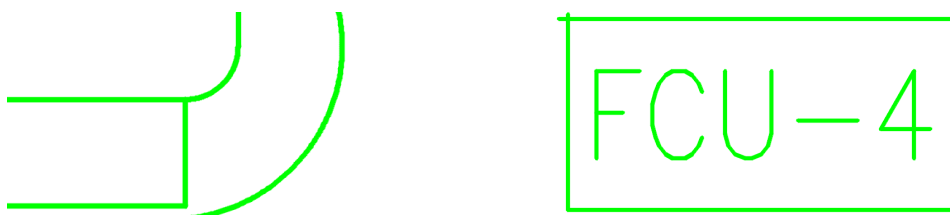
(a) The line segment is not closed correctly.   (b) Line segment out of range.

712   **Fig. 20.** Examples of drafting errors.

713   ## 8. Conclusions

714   Recovering BIM from 2D drawings to support the effective O&M of MEP
715   systems helps conserve energy and reduce emissions in ageing buildings. However,
716   the precise semantic and geometric information needed for O&M is difficult to obtain
717   from the drawings, primarily due to the multiscale and irregular nature of MEP
718   components. This study aimed to recover information-rich BIM with accurate shapes
719   from 2D MEP drawings to facilitate the BIM application in the O&M of MEP systems
720   in ageing buildings. To this end, a three-module framework for automatic MEP model
721   reconstruction was developed. Compared to previous studies, the framework features
722   the following innovations.

723   The objective of the semantic information extraction module was to extract
724   accurate semantics and coarse contours of MEP components. Previous studies [23,30]
725   used handcrafted features to obtain semantic information and represented MEP
726   components using rectangular boxes. As such, the accuracy of the extracted semantics
727   and geometry is prone to being affected by the diverse, multiscale, and irregular
728   components. In contrast, the developed module adopts the Cascade Mask R-CNN
729   model, which can detect diverse and irregular components robustly. Besides, when
730   combining layered drawings with the image cropping strategy, the module enables
731   multiscale object segmentation in large-size raster drawings. Experimental results
732   show that the model has a mAP over 0.9 in both the object detection and semantic
733   segmentation tasks, indicating that the module can accurately recognise the semantics
734   and contours of multiscale irregular components.

735   The geometry extraction module was developed to extract components' fine
736   geometry. Existing studies either use pre-defined rules [22] or the combination of
737   object detection with annotations [17] for geometric extraction. However, these
738   practices are unsuitable for irregular MEP components, whose amorphous shapes
739   cannot be precisely described by rules or annotations. Unlike these studies, the
740   developed module takes advantage of semantic information and extracts the shape and
741   location of components through semantic-assisted image processing. The robust
742   semantics of instance segmentation help reduce the interference from other irrelevant
743   components in geometric extraction. The results show that the error in the area and
744   perimeter of the contours detected is less than 5%, which validates the accuracy and
745   reliability of this module.

746   The BIM automatic reconstruction module was designed to produce IFC for MEP
747   components. Most IFC schema-based studies' modelling processes are complicated
748   and inefficient due to a lack of support for modelling irregular-shaped components.
749   Compared to these studies, the developed module leverages pythonOCC and
750   IfcOpenShell to realise automated IFC modelling for irregular MEP components. The
751   module takes only 2.85s and 0.79s to generate two MEP systems containing 335 and

282 regular and irregular components, respectively. These results demonstrate the superior modelling efficiency of this module. In conclusion, the developed framework improves the efficiency and accuracy of MEP model reconstruction. It will provide reliable semantic and geometric data on MEP components to aid in the O&M of MEP systems in ageing buildings.

Acknowledgements

References

[1] D.A. Waddicor, E. Fuentes, L. Sisó, J. Salom, B. Favre, C. Jiménez, M. Azar, Climate change and building ageing impact on building energy performance and mitigation measures application: A case study in Turin, northern Italy, Building and Environment. 102 (2016) pp. 13–25. https://doi.org/10.1016/j.buildenv.2016.03.003.

[2] A. Hamburg, T. Kalamees, How well are energy performance objectives being achieved in renovated apartment buildings in Estonia?, Energy and Buildings. 199 (2019) pp. 332–341. https://doi.org/10.1016/j.enbuild.2019.07.006.

[3] D. Lee, C.-C. Cheng, Energy savings by energy management systems: A review, Renewable and Sustainable Energy Reviews. 56 (2016) pp. 760–777. https://doi.org/10.1016/j.rser.2015.11.067.

[4] P. de Wilde, Ten questions concerning building performance analysis, Building and Environment. 153 (2019) pp. 110–117. https://doi.org/10.1016/j.buildenv.2019.02.019.

[5] T.-W. Kang, H.-S. Choi, BIM perspective definition metadata for interworking facility management data, Advanced Engineering Informatics. 29 (2015) pp. 958–970. https://doi.org/10.1016/j.aei.2015.09.004.

[6] D. Utkucu, H. Sözer, Interoperability and data exchange within BIM platform to evaluate building energy performance and indoor comfort, Automation in Construction. 116 (2020) pp. 103225. https://doi.org/10.1016/j.autcon.2020.103225.

[7] L.-T. Wu, J.-R. Lin, S. Leng, J.-L. Li, Z.-Z. Hu, Rule-based information extraction for mechanical-electrical-plumbing-specific semantic web, Automation in Construction. 135 (2022) pp. 104108. https://doi.org/10.1016/j.autcon.2021.104108.

[8] Y.-Q. Xiao, S.-W. Li, Z.-Z. Hu, Automatically generating a MEP logic chain from building information models with identification rules, Applied Sciences. 9 (2019). https://doi.org/10.3390/app9112204.

791 [9]   Z.-Z. Hu, P.-L. Tian, S.-W. Li, J.-P. Zhang, BIM-based integrated delivery
792   technologies for intelligent MEP management in the operation and
793   maintenance phase, Advances in Engineering Software. 115 (2018) pp. 1–16.
794   https://doi.org/10.1016/j.advengsoft.2017.08.007.

795 [10] R. Volk, J. Stengel, F. Schultmann, Building Information Modeling (BIM) for
796   existing buildings — Literature review and future needs, Automation in
797   Construction.   38   (2014)   pp.   109–127.
798   https://doi.org/10.1016/j.autcon.2013.10.023.

799 [11] Q. Wang, J. Guo, M.-K. Kim, An application oriented Scan-to-BIM framework,
800   Remote Sensing. 11 (2019). https://doi.org/10.3390/rs11030365.

801 [12] F. Bosche, C.T. Haas, Automated retrieval of 3D CAD model objects in
802   construction range images, Automation in Construction. 17 (2008) pp. 499–512.
803   https://doi.org/10.1016/j.autcon.2007.09.001.

804 [13] C. Zhang, D. Arditi, Automated progress control using laser scanning
805   technology, Automation in Construction. 36 (2013) pp. 108–116.
806   https://doi.org/10.1016/j.autcon.2013.08.012.

807 [14] B. Wang, Q. Wang, J.C.P. Cheng, C. Song, C. Yin, Vision-assisted BIM
808   reconstruction from 3D LiDAR point clouds for MEP scenes, Automation in
809   Construction.   133   (2022)   pp.   103997.
810   https://doi.org/10.1016/j.autcon.2021.103997.

811 [15] Y. Kim, C.H.P. Nguyen, Y. Choi, Automatic pipe and elbow recognition from
812   three-dimensional point cloud model of industrial plant piping system using
813   convolutional neural network-based primitive classification, Automation in
814   Construction.   116   (2020)   pp.   103236.
815   https://doi.org/10.1016/j.autcon.2020.103236.

816 [16] J. Zhang, IFC-based information extraction and analysis of HVAC objects to
817   support building energy modeling, Proceedings of the 39th International
818   Symposium on Automation and Robotics in Construction. (2022) pp. 159–166.
819   https://doi.org/10.22260/ISARC2022/0024.

820 [17] Y. Zhao, X. Deng, H. Lai, Reconstructing BIM from 2D structural drawings for
821   existing buildings, Automation in Construction. 128 (2021) pp. 103750.
822   https://doi.org/10.1016/j.autcon.2021.103750.

823 [18] Z. Ma, S. Liu, A review of 3D reconstruction techniques in civil engineering
824   and their applications, Advanced Engineering Informatics. 37 (2018) pp. 163–
825   174. https://doi.org/10.1016/j.aei.2018.05.005.

826 [19] P. Tang, D. Huber, B. Akinci, R. Lipman, A. Lytle, Automatic reconstruction
827   of as-built building information models from laser-scanned point clouds: A
828   review of related techniques, Automation in Construction. 19 (2010) pp. 829–
829   843. https://doi.org/10.1016/j.autcon.2010.06.007.

830 [20] C.Y. Cho, X. Liu, An automated reconstruction approach of mechanical
831   systems in Building Information Modeling (BIM) using 2D drawings, in:
832   Computing   in   Civil   Engineering   2017,   2017:   pp.   236–244.

833   https://doi.org/10.1061/9780784480823.029.

834   [21] L. Gimenez, S. Robert, F. Suard, K. Zreik, Automatic reconstruction of 3D
835   building models from scanned 2D floor plans, Automation in Construction. 63
836   (2016) pp. 48–56. https://doi.org/10.1016/j.autcon.2015.12.008.

837   [22] Q. Lu, L. Chen, S. Li, M. Pitt, Semi-automatic geometric digital twinning for
838   existing buildings based on images and CAD drawings, Automation in
839   Construction. 115 (2020) pp. 103183.
840   https://doi.org/10.1016/j.autcon.2020.103183.

841   [23] C.Y. Cho, X. Liu, B. Akinci, Automated building information models
842   reconstruction using 2D mechanical drawings, in: I. Mutis, T. Hartmann (Eds.),
843   Advances in Informatics and Computing in Civil and Construction Engineering,
844   Springer International Publishing, Cham, 2019: pp. 505–512.
845   https://doi.org/10.1007/978-3-030-00220-6_60.

846   [24] M. Yin, L. Tang, T. Zhou, Y. Wen, R. Xu, W. Deng, Automatic layer
847   classification method-based elevation recognition in architectural drawings for
848   reconstruction of 3D BIM models, Automation in Construction. 113 (2020) pp.
849   103082. https://doi.org/10.1016/j.autcon.2020.103082.

850   [25] Z. Fan, L. Zhu, H. Li, X. Chen, S. Zhu, P. Tan, FloorPlanCAD: A large-scale
851   CAD drawing dataset for panoptic symbol spotting, in: 2021 IEEE/CVF
852   International Conference on Computer Vision (ICCV), 2021: pp. 10108–10117.
853   https://doi.org/10.1109/ICCV48922.2021.00997.

854   [26] L.-P. de las Heras, S. Ahmed, M. Liwicki, E. Valveny, G. Sánchez, Statistical
855   segmentation and structural recognition for floor plan interpretation,
856   International Journal on Document Analysis and Recognition (IJDAR). 17
857   (2014) pp. 221–237. https://doi.org/10.1007/s10032-013-0215-2.

858   [27] S.K. Maity, B. Seraogi, S. Das, P. Banerjee, H. Majumder, S. Mukkamala, R.
859   Roy, B.B. Chaudhuri, An approach for detecting circular callouts in
860   architectural, engineering and constructional drawing documents, in: A. Fornés,
861   B. Lamiroy (Eds.), Graphics Recognition. Current Trends and Evolutions,
862   Springer International Publishing, Cham, 2018: pp. 17–29.
863   https://doi.org/10.1007/978-3-030-02284-6_2.

864   [28] S. Macé, H. Locteau, E. Valveny, S. Tabbone, A system to detect rooms in
865   architectural floor plan images, in: Proceedings of the 9th IAPR International
866   Workshop on Document Analysis Systems, Association for Computing
867   Machinery, New York, NY, USA, 2010: pp. 167–174.
868   https://doi.org/10.1145/1815330.1815352.

869   [29] M.C. Cooper, Formal hierarchical object models for fast template matching,
870   The Computer Journal. 32 (1989) pp. 351–361.
871   https://doi.org/10.1093/comjnl/32.4.351.

872   [30] C.Y. Cho, X. Liu, B. Akinci, Symbol recognition using vectorial signature
873   matching for building mechanical drawings, Advances in Computational
874   Design. 4 (2019) pp. 155–177. https://doi.org/10.12989/ACD.2019.4.2.155.

[31] Y. Zhao, X. Deng, H. Lai, A deep learning-based method to detect components from scanned structural drawings for reconstructing 3D models, Applied Sciences. 10 (2020). https://doi.org/10.3390/app10062066.

[32] Y. Xiao, S. Chen, Y. Ikeda, K. Hotta, Automatic recognition and segmentation of architectural elements from 2D drawings by convolutional neural network, in: D. Holzer, W. Nakapan, A. Globa, I. Koh (Eds.), The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), The Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), 2020: pp. 843–852. https://doi.org/10.52842/conf.caadria.2020.1.843.

[33] Y. Wu, J. Shang, P. Chen, S. Zlatanova, X. Hu, Z. Zhou, Indoor mapping and modeling by parsing floor plan images, International Journal of Geographical Information Science. 35 (2021) pp. 1205–1231. https://doi.org/10.1080/13658816.2020.1781130.

[34] R. Vanlande, C. Nicolle, C. Cruz, IFC and building lifecycle management, Automation in Construction. 18 (2008) pp. 70–78. https://doi.org/10.1016/j.autcon.2008.05.001.

[35] A. Van Etten, Satellite imagery multiscale rapid detection with windowed networks, in: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), 2019: pp. 735–743. https://doi.org/10.1109/WACV.2019.00083.

[36] J. Wang, H. Xiao, L. Chen, J. Xing, Z. Pan, R. Luo, X. Cai, Integrating weighted feature fusion and the spatial attention module with convolutional neural networks for automatic aircraft detection from SAR images, Remote Sensing. 13 (2021). https://doi.org/10.3390/rs13050910.

[37] Z. Cai, N. Vasconcelos, Cascade R-CNN: High quality object detection and instance segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence. 43 (2021) pp. 1483–1498. https://doi.org/10.1109/TPAMI.2019.2956516.

917
918
919
920
921
922
923
924