IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 17 October 2023, accepted 12 November 2023, date of publication 27 November 2023, date of current version 4 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3337049

SURVEY

Point Cloud Analysis of Railway Infrastructure: A Systematic Literature Review

BRAM DEKKER[1], BRAM TON[1,2], JOANNEKE MEIJER[2], NACIR BOUALI[1], JEROEN LINSSEN[2], AND FAIZAN AHMED[1,2]
[1]Department of Computer Science, University of Twente, 7522 NB Enschede, The Netherlands
[2]Ambient Intelligence Group, Saxion University of Applied Sciences, 7513 AB Enschede, The Netherlands

Corresponding author: Faizan Ahmed (f.ahmed@utwente.nl)

This work was supported in part by the Tech For Future Grant 2207 ''Digital Twinning voor Spoorontwerp'' and Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) under Grant NWA.1160.18.238, and in part by the University of Twente and the Saxion University of Applied Sciences.

ABSTRACT Digitalisation in railway networks harnesses digital technologies to optimise operations, leading to enhanced efficiency, and reduced energy consumption. By analysing real-time data, railways can predict maintenance needs, improve passenger experiences, and seamlessly integrate with other transport modes. As societies strive for sustainable transportation solutions, it is imperative to understand and collect digitalisation techniques to enhance efficiency and reduce the ecological footprint of railway networks. This paper serves as a snapshot of the current state of the art addressing the pivotal role of point cloud techniques in advancing railway digitalisation and providing valuable pointers for future research directions. Employing a systematic review approach, our study concentrates exclusively on research centred around railway assets and their digitalisation via point cloud data. We have themed the literature into pre-processing, modelling, and digital twinning. Within this review, we analyse diverse modelling and pre-processing techniques and categorise them for clarity. The digital twin techniques are also collected, though these techniques are scarce in the context of railway infrastructure and point clouds. The paper also presents a compilation of dataset statistics highlighting the scarcity of openly available railway-specific datasets. This scarcity considerably hampers the feasibility of research reproducibility and the comparative analysis of different approaches. Our conclusion reflects on the challenges encountered and proposes a course for future research. Particularly, we conclude that hybrid methodologies that combine machine learning with structure-based techniques hold substantial promise toward creating digital twins, considering the intrinsic characteristics of railway infrastructure.

INDEX TERMS Railways, point clouds, digitalization, infrastructure, deep learning, digital twin.

I. INTRODUCTION

Compared to other means of transport, such as air or road transportation, rails are viewed as being more environment-friendly [1]. To maximise the benefits of this greener and sustainable alternative, initiatives such as SHIFT2RAIL and European Rail Traffic Management System (ERTMS) have been initiated by the EU [1], [2]. The latter aims to digitalise the management of railway infrastructure to increase its capacity and lessen its greenhouse gas emissions. The ERTMS digitalisation efforts aim for safer, more competitive, and more integrated railway infrastructure [1, Page 23] also [3]. One of the objectives of ERTMS is to further improve the predictive maintenance in railway infrastructure through improved early fault detection [1].

Any railway digitalisation effort is met with context-sensitive challenges mainly related to the criticality of the railway system, its legacy systems, interoperability, data integration and standardisation, cyber-security, reliability and safety, regularity and organisational challenges [4]. We focus in the rest of this review on the digitalisation of the railway infrastructure, in which imaging technologies, such

The associate editor coordinating the review of this manuscript and approving it for publication was Jesus Felez.

as Light Detection And Ranging (LiDAR), play a crucial role. Other sensors, simpler in this context, such as cameras provide visual information, however, their use in the railway infrastructure is both limited and limiting due to the required lighting conditions, depth perception, and privacy.

Point clouds are captured using LiDAR. This is done through laser scanning, which operates by emitting a usually non-visible laser light pulse and calculating its time-of-flight. This time is directly proportional to the distance of the object. Three main configurations of laser scanning exist:

- *Terrestrial Laser Scanner (TLS):* The laser scanner is usually fixed on a tripod.
- *Mobile Laser Scanner (MLS):* The laser scanner is mounted on a mobile carrier platform such as a car, boat or train.
- *Airborne or Aerial Laser Scanner (ALS):* The laser scanner is attached to an airborne carrier platform such as an unmanned aerial vehicle (UAV), helicopter or airplane.

LiDAR is reputed for precise mapping, extensive range capabilities, its ability to see through vegetation, and compatibility with other tech systems. Its applications span sectors like agriculture, environmental monitoring, archaeology, and forestry [5]. The output of the LiDAR is recorded in the form of point clouds. A *point cloud* $\mathcal{P}$ is a finite set of points with cardinality n in $\mathbb{R}^3$. Associated with each individual point there is an *optional* feature vector $\mathbb{F}^D$. This D dimensional feature vector can contain information such as reflection intensity or colour information. Point clouds could be used to create a digital model of the railway infrastructure that can act as a digital twin. This twin, when regularly updated, can serve for continuous monitoring.

The conversion of point clouds into a digital twin is a complex, multi-step process. An initial step often involves segmenting the cloud into different railway-related objects, such as tracks or poles. The interest in point cloud segmentation is not exclusive to railway infrastructure monitoring, it is rather crucial for other domains, especially in autonomous driving [6] or infrastructure monitoring [7] among many other applications [8].

LiDAR technology, while not new, has gained renewed interest due to the surge of machine learning-based techniques [9]. The research on the crosscut between LiDAR and machine learning is multi-faceted, emphasising the need to collate and analyse literature on point clouds within the railway monitoring and predictive maintenance domain.

Previous systematic reviews have addressed 3D data collection and analysis, railway datasets, and point cloud analysis methods. For instance, [10] discusses data integration of different domains to obtain a 3D dataset of the railway environment. Dong et al. reviews methods for the registration of terrestrial laser scanner point clouds [11]. Different datasets of the railway environment are discussed in [12]. Techniques for point cloud analysis are reviewed in [9] and [13]. However, there seems to be a gap in systematic reviews specifically targeting point cloud segmentation or object detection methods.

This ***review aims*** to provide an overview of the current state-of-the-art methods, models, and technologies that can be used to digitalise railway infrastructure for monitoring and maintenance.

Railway scene, railway environment, and railway infrastructure are all closely related terms with similar meanings. To avoid ambiguity, we list the definitions below as used in this research:

- *Railway scene:* All objects in the surroundings of the railway tracks including vegetation, urban buildings and foreign objects.
- *Railway environment:* Synonym for railway scene.
- *Railway infrastructure:* All objects specifically belonging to the railway like tracks, poles, catenary arches, wires etc. These are the objects of interest for this study.

The remainder of this review is structured as follows: Section II describes the review strategy. Section III provides metadata about the publications and includes a table summarising the characteristics of the datasets used in the included studies. The paper focuses on the gathering of literature for pre-processing (Section IV), modelling (Section V), and the creation of a digital twin (Section VI). The discussion section (Section VII) reflects on the gathered literature, identifies the literature gap, and provides future directions. Section VIII concludes the review.

II. REVIEW METHOD

In this section, the research method used to conduct this literature review is presented. We have used Covidence to manage the review process. Covidence is a web-based collaboration software platform that streamlines the production of systematic and other literature reviews [14].

A. REVIEW QUESTION

The main research question for this systematic literature review is:

What is the state of the art in point cloud analysis (both classification and segmentation) of railway infrastructure?

B. DATA SOURCES AND SEARCH STRATEGY

To select proper data sources to find articles for our literature review, the following criteria are used:

include only databases that are pertinent to our research (only general databases, engineering databases or computer science specific databases)

include only databases that have peer-reviewed articles

include only databases that allow to search on phrases

To extract paper relevant to the research question we have primarily used two databases:

- Scopus [15]
- Web of Science [16]

We have also used three other databases to verify the completeness of information namely ACM [17], DBLP [18], and IEEExplore [19].
The search query used for finding relevant literature was:

(point cloud OR point clouds) AND railway

We restricted the search to the papers' titles, abstracts, and keywords with a case-insensitive search. In Scopus, a total of 271 papers were found, and in Web of Science, 158 papers were found. Importing all these papers in Covidence resulted in 121 duplicates, thus leaving 308 papers for further review. We have manually checked the results from the other databases and compared them with the list generated in Covidence. This comparison did not reveal any new papers.

C. STUDY SELECTION

The papers were further screened by reading the title and abstract. Only papers satisfying all criteria proceeded to a full-text review, the others were excluded. The inclusion criteria used in this study are:

1) **include** papers written in English or Dutch
2) **include** papers published in 2005 or later
3) **include** papers using outdoor data
4) **include** papers describing methods of analysing/pre-processing point clouds
5) **include** papers describing scenery reconstruction if the dataset contains tunnels/bridges
6) **include** only papers describing transformation from point cloud to mesh
7) **include** only papers describing Building Information Modelling of railway infrastructure

Criterion 5 was included after the observation during the abstract screening phase that some point cloud papers are only handling tunnels and bridges. The railway infrastructure was not specifically included. However, some examples contained parts of railways. This is observed mainly for papers focused on the deformation of railway tunnels, where the focus was on the tunnel structure instead of railway infrastructure such as railway lines or catenary arches.
Also some exclusion criteria were used:

exclude papers only describing geometry in point clouds

exclude papers only describing foreign object detection on the rail tracks

exclude short papers (less than four pages long)

Shorter papers, often less than four pages, may lack the comprehensive details and thoroughness found in longer articles, potentially offering only preliminary findings or lacking in-depth methodologies. Such papers might not have undergone the same rigorous peer review process as full-length articles, which is a vital step in ensuring the validity and quality of research. Therefore, to maintain the integrity and depth of our review, we have chosen to exclude papers less than four pages.

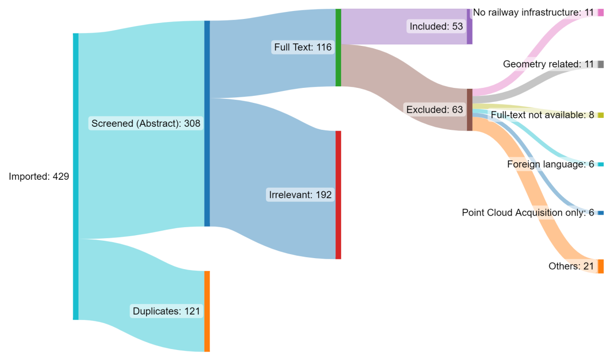

FIGURE 1. An overview of the paper selection process with exclusion criteria.

1) SCREENING PROCESS

At the abstract and title screening stage, at least two assessors screened each paper. If the assessors disagreed on including or excluding the paper, a third assessor screened the title and abstract and decided the outcome. The procedure is applied to all 308 papers and resulted in the exclusion of 192 papers. Thus, 116 papers are left for the full-text review stage.

A single assessor conducted the full-text review for each paper. Should there be any doubt regarding discarding a paper, it was referred to a second assessor. The rationale for a paper's rejection is duly recorded. Following this method, another 63 papers were excluded from the data extraction phase. This leaves 53 papers relevant to our research question which proceeded to the data extraction phase. Figure 1 summarises the screening process and also detailing the number of papers excluded at each stage.

D. DATA EXTRACTION

In order to maintain the consistency of the data extraction process, we have used a form. The form is given in Table 1. For the digitalisation of infrastructure, data collection plays a crucial role. Therefore, we have collected data reported in the selected studies related to data collection or meta-information. Most importantly, we collected the scan speed (speed of the vehicle if it is vehicle mounted), presence of colour information, or simultaneous collection of other sensory data such as GPS. Note that not all papers have described the data collection process.

We divided infrastructure digitalisation into three stages. The first stage is pre-processing, where raw or filtered data is pre-processed for modelling purposes. The second stage is the modelling itself, while the last stage is the creation of digital twins. The 'Steps' field is used to register which stages are described in the paper. The results section is also divided with respect to these stages. The notes field is used to note down any other relevant information not covered by any other field.

III. META ANALYSIS, CHALLENGES, AND DATASETS

To get a better insight into the gathered data, clusters of articles are formed based on common characteristics like

TABLE 1. The data extraction form used for gathering useful information from every article.

Field name	Description
Title	The title of the paper
Goal	The main goal of the paper
Steps	Which steps are described in the paper (pre-processing, modelling, digital twinning)?
Pre-processing techniques	List the used pre-processing techniques
Modelling techniques	List the used modelling techniques
Digital twinning techniques	List the used digital twinning techniques
Measurement of quality	What quality metrics are used?
Country	From which country is the data in the dataset?
Dataset	List characteristics of the dataset
Experimental setup: speed	Speed of the vehicle while collecting the dataset
Experimental setup: scanner type	Type of scanner used for collecting the dataset
Experimental setup: RGB data	Are colours (or other additional information) present in the dataset?
Notes	Any other relevant and interesting aspects

publication year, nature of the dataset or analysis method used.

It is apparent from Figure 2 that the point cloud analysis for railway scenes has been gaining interest in recent years. Another interesting observation is the dip in the number of publications for 2017-2019, with a further increase in 2020/2021. We cannot associate a reason to the dip in the number of publications. However, the increase can be attributed to the popularity of deep learning-based techniques. The rise in the utilization of deep learning techniques for point cloud data analysis can be significantly attributed to the seminal paper "PointNet" by Qi et al. in 2017 [20]. This work was groundbreaking because it introduced a novel neural network that could process point clouds directly. Note that the data for the year 2022 is incomplete since the query was run in November 2022.

It was evident from the full-text search that most papers can be categorised into three classes based on the objective of the analysis. These steps were pre-processing, modelling, and digital twin. All papers have at least one of these aspects as the main contribution. The distribution of papers according to steps is given in Figure 3. It is clear from the table that there is no single paper with digital twins as a core focus. In most cases, it is combined with modelling. A combination of pre-processing and modelling is understandably the most used.

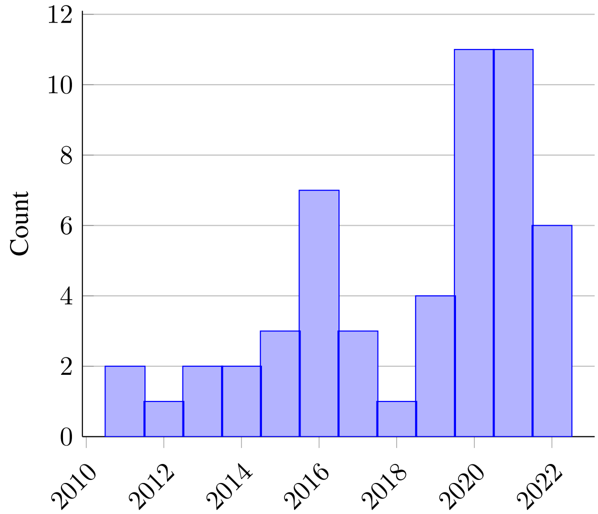

FIGURE 2. Number of publications per year (from papers included in this study).

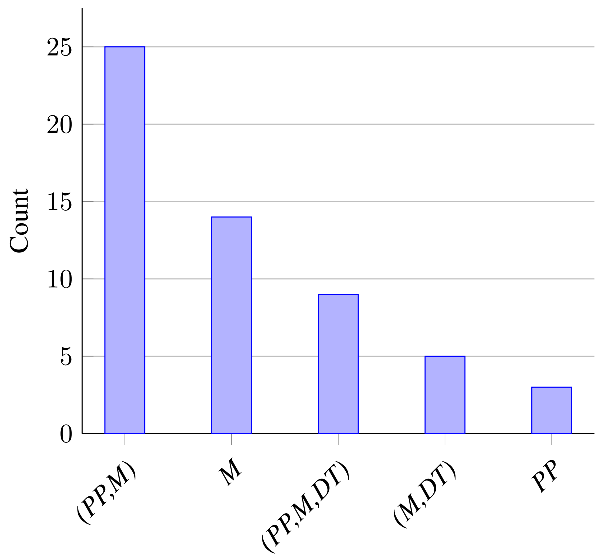

FIGURE 3. An overview of the count of publication describing pre-processing(PP),modelling(M), and digital twin (DT). Note that there is no paper with sole focus on digital twin.

A. DATASET COLLECTION AND BENCHMARK DATASET

One interesting finding of our literature review is the lack of public benchmark datasets consisting of point clouds in the context of railway infrastructure. However, we have recently published a fully labelled dataset consisting of catenary arches [21], which is the only openly available dataset to the best of our knowledge. Although a few datasets are mentioned in the literature, they are not openly accessible. The only paper we found concerning data collection in the context of railway infrastructure is [22] that have reported the most detailed data collection methodology. The authors have presented the approach together with pre-processing. The primary focus was on change detection for the safety and security of railway infrastructure [22].

The datasets from the included studies are summarised in Table 2. The table presents a total of 46 datasets from diverse geographical locations, predominantly from China (11), the European Union (24), and other countries (11), showcasing global research interest. Various data acquisition methods are employed across studies. The majority of the studies used MLS (31), followed by ALS (8), TLS (3) and other methods (4). The datasets vary largely in terms of point density, ranging from densities as low as 50 points/m^2 to as high as 2,500 points/m^2, and cover short stretches (80 m) to several kilometres (120 km). Additionally, while many studies focus on geometric data, only the minority of the datasets incorporate RGB information, highlighting the multifaceted nature of the research.

In the process of collating data for the table, we occasionally derived the density or length values from other information provided within the papers. A notable observation was the complete absence of publicly available datasets. While many papers emphasised the significance of point density, it was interesting to see that a quantitative report on density was often omitted rather it was described qualitatively like low or high density. Interestingly, there was a dataset that focused on lab-generated data of bolts [23], but we chose to exclude it from the table for clarity. A particularly remarkable dataset [24], originated from China. Despite being recorded at an impressive speed of 193 km/h, it boasted an exceptionally high point density of 3000 points/m^2, underscoring the advancements in data acquisition techniques. For some datasets, we assumed that they were the same because they are from the same research group and have the same characteristics, although it was not stated explicitly in the papers.

B. CHALLENGES OF POINT CLOUD DATA

Point clouds are irregular, unstructured and unordered, unlike 2D images, and are thus a challenging data type to work with [72]. Following is a list of the most significant challenging characteristics that are inherent to point cloud data. Sensor type, environment, weather conditions and sensing distance influence the degree to which point clouds suffer from these characteristics [6]:

- **Irregularity**: point clouds usually have non-uniform distributed point density.
- **Unstructured**: point clouds are not placed on a regular grid. Each point is scanned independently, and its distance to neighbouring points is not fixed. This also means that voxelisation of point clouds often leads to empty voxels, i.e. data sparsity.
- **Unordered**: a point cloud is a set of points, the order in which the points are stored does not change the representation.
- **Size**: point clouds often contain millions of points taking up large chunks of memory and thus it is time-consuming to process and analyse them.

TABLE 2. An overview of the datasets used in the included papers.

Ref.	Country	Scanner type	Speed (km/h)	Density (p/m^2)	Length (m)	RGB
[25], [26]	Netherlands	ALS	75	-	80	No
[25]	Netherlands	TLS	-	-	630	No
[27]	Austria	MLS	125	-	550	Yes
[28], [29]	Netherlands	ALS	-	293	18.000	No
[30]	Austria	ALS	-	60-90	120.000	Yes
[31]–[33]	Germany	TLS	-	-	500	No
[34]	Germany	MLS			50.000	No
[35]	China	TLS	-	-	885	No
[36]	China	MLS	3.6	-	16.700	No
[37]	China	MLS	-	-	150	No
[38]	Italy	ALS	-	-	1000	No
[39]	Hungary	MLS	60	-	34.000	Yes
[40]	China	-	-	-	2000	No
[41]	Italy	Stereo vision	-	-	300	Yes
[42]	China	ALS	-	-	-	No
[43]	Portugal	MLS	-	-	550	No
[44]–[46]	Spain	MLS	10	980	90.000	No
[47]	Spain	MLS	4	2500	-	No
[48]	South Korea	MLS	50-70	100-800	1000	No
[49]	South Korea	MLS	60	85	100	No
[50]	Germany	MLS	-	-	-	No
[51]	Japan	MLS	20	1600	-	No
[52]	Poland	ALS	-	10-17	772	No
[52]	Poland	MLS	-	-	550	No
[53]	China	MLS	120	-	3500	No
[54]–[56]	China	MLS	3.6	1028	16.000	No
[57]	France	MLS	-	-	-	Yes
[58]	France	ALS	-	-	5250	No
[58]	France	MLS	-	-	13.000	No
[59]	Taiwan	MLS	-	-	14.000	No
[60]	Austria	MLS	-	700	600	No
[61]	Poland	MLS	-	-	90.000	No
[34]	Germany	MLS	-	-	51.000	No
[62]	England	MLS	-	-	16.100	No
[63]	Iran	Aerial images	-	-	200	Yes
[64]	Europe	-	-	-	90.000	No
[22]	Italy	MLS	0.36-3.6	-	-	Yes
[65]	Sweden	MLS	-	568	2000	No
[66]	Hong Kong	MLS	-	-	16.000	No
[67]	China	MLS	60	-	100.000	No
[68]	USA	MLS	-	77	2000	No
[69]	China	MLS	-	-	-	No
[24]	China	MLS	193	3000	11.690	No
[70]	Finland	MLS	35	720	2000	No
[70]	Finland	ALS	-	50	2000	No
[71]	China	MLS	-	490	-	No

- **Measurement artefacts**: point clouds can contain noise in the data produced for example by errors of the scanner or moving objects [73].
- **(Partial) Occlusion**: point clouds suffer from (partial) occlusion of objects since other objects may block them [74].

A challenge for railway scenes is the large variance in object sizes (a top bar can be well over 20 metres long, while an insulator typically is around 30 centimetres [21], which is a size ratio of at least 60 times). An additional challenge is the huge class imbalance encountered within the rail environment, for instance certain objects like masts occur very regularly, but relay cabinets occur a lot less often.

IV. PRE-PROCESSING

Point clouds are unordered sets of points. Absence of structure makes them a challenging datatype to deal with. Pre-processing techniques help to reduce the volume of data, introduce a structure or filter out the dispensable points. In certain cases the boundary between pre-processing and modelling is blurred due to the fact that the result of pre-processing is sometimes already a feature. Therefore, we do not apply the term in their strict sense instead focus on the mechanisms of the techniques.

In general, the main goal of the pre-processing is to cull points such that further processing steps require less computational effort. In the following subsections we list the pre-processing techniques found in literature and their associated references.

A. CROPPING

Cropping is a very rudimentary pre-processing step that removes points based on a specified bounding region. This is predicated on the assumption that the points outside this region do not contain information of interest. For instance, the work of Ariyachandra and Brilakis, which focuses on detecting elements of the overhead line equipment, remove all points belonging to the ground by setting a threshold value of 0.23 cm. Points with a z-coordinate below this threshold are removed [29]. Similarly Chen et al. also use fixed thresholds to remove distant points with no information [36]. A more advanced method of detecting ground points is proposed by Chen et al. which use a Euclidean distance clustering segmentation algorithm [37]. When point clouds are collected using a mobile scanner mounted on a train, the trajectory log can play an important role in the culling of points. As an example, Pastucha defines an extent of 5 m on both sides of the trajectory. Points outside of this region are removed. The scan angle, which is usually recorded as meta-data of a point, can also be used as a filter condition to remove points [43]. As an example of how the scan angle can be used to crop relevant regions of points, the authors show how the track centre lines and the ballast top can easily be recognised from the point cloud data. To remove vegetation, the work of Cserép et al. first project the scene to 2D by registering the maximum value of the z-coordinate. After this contour detection is used to filter out vegetation [39], unfortunately no further details are provided for this approach.

Which points to cull is also highly dependent on the application. If the application is to detect tracks, it makes sense to only maintain points which relate to the tracks. Specifically for this purpose, Ponciano et al. use a mask-based approach to only keep points which relate to the tracks [34]. An alternative approach provided by Zou et al. first filter the point cloud based on intensity values, only values with a low intensity are kept. After filtering, tracks remain, but still there is significant noise. Further refinement steps are required to extract the tracks [71].

B. PARTITIONING

Commonly the point cloud data provided covers a large area. In order to create tractable pieces that can be used in downstream processing steps the larger point cloud is usually partitioned into smaller pieces. Ariyachandra and Brilakis manually partitioned a large point cloud that covered ≈18 km into three pieces covering ≈6 km each [29]. In a related work, the same authors employed an optimisation strategy to determine the optimal number of partitions for splitting the dataset [28]. Constraints used in this optimisation approach were the curvature of the track, number of noise points, and the cropping of masts. The width of the scenes was limited to 30 m.

The work of Lamas et al. use the trajectory log of the measurement train to partition the data into pieces which are 100 m long and 20 m wide [45]. Pastucha uses even smaller sections which are 0.5 m in length [61]. Surprisingly, only a limited number of studies utilise the raw frame-by-frame data from scanner, with most relying solely on aggregated results. An exception is the work of Chen et al. that use 2D laser scan lines to segment the overhead contact system [36]. This raw frame data is commonly used for applications such as autonomous driving. The envisioned benefit of using this raw data is that the data will have a fixed frame of reference, i.e. it is always known how the data is captured with reference to the current track.

C. NORMALISATION

Normalisation of the training data plays an important role, especially when deep learning methods are involved. To align individual pieces of point cloud data along the x-axis Ariyachandra and Brilakis use a Principle Component Analysis (PCA) to determine the major axis of the point cloud [28]. The work of Lamas et al. also use PCA, albeit in a slightly modified form, to align the direction of the tracks along the x-axis. Corongiu et al. align the point cloud subsets to the y-axis, unfortunately the method to do so is not described [38]. The trajectory log of the mobile sensing platform facilitates a convenient way of aligning sub-point cloud to the track [61]. Of course, the aforementioned partitioning of the scene into regular-sized pieces is also a form of normalisation.

D. PROJECTION

As point cloud data has no structure, sometimes the point cloud is projected to a 2D plane with a grid to create an image. This image can then be processed with conventional image processing techniques. For example, Corongiu et al. flatten the point cloud to a 2D grid by summing in the z-direction. Within this image masts will be visible as high-intensity blobs, making it easy to locate them [38].

An interesting piece of work, albeit in a very premature state, is presented by Wolf et al. Their approach to detect railway assets from point cloud data is to first render a greyscale image from a slice of point cloud data [75]. The pixel values are the intensity values from the original point cloud data. These slices are taken perpendicular to the rail track. The work shows results of both object detection, based on the YOLOv3 model [76], and on semantic segmentation, based on U-Net [77]. An image-based approach has two major benefits: the field of image processing has advanced much further than point-based methods and the processing of raster data can be done much more efficiently compared to point data.

E. DATA STRUCTURES

Voxelisation is the process of defining a regular 3D grid, each element of the grid is referred to as a voxel. This is analogous to a pixel in the 2D case. The benefit of the voxelisation process is that it creates a structured format which can be processed very efficiently. For example, Jung et al. extract line segments per voxel [48]. Another data structure which occurs is the *kd*-tree, this data structure is used for efficiently selecting neighbour points around a query point [27].

When point clouds are captured using a laser scanner, the captured point density close to the sensor is higher compared to regions further away from the sensor. To homogenise the density across the entire scene, a fixed number of points per voxel can be retained [44], [45]. Not only does this improve the homogeneity of the point distribution, but it also reduces the number of points.

Besides voxelisation, different grid definition schemes are possible. For instance, Yu et al. use pyramid partitions [69]. This approach defines smaller volumes close to the sensor and increases the volume gradually when the distance to the sensor increases. This ensures that the number of points per volume remains roughly the same.

F. SAMPLING

Down-sampling is a common pre-processing step to reduce the number of points or to achieve a fixed number of points [40], [44]. Fixed number of points are usually required when training deep learning models. For instance, Grandio, Riveiro, Soilán, et al. used a fixed size of 16384 (2^{14} and 32768 (2^{15}) points for training a PointNet++ segmentation model [44]. Note that it is a common misconception that such models *require* a fixed number of points as input. The architecture of these models are agnostic of the point set size, but the frameworks used to implement the models are the bottleneck.

To create tractable pieces which can be used during training of a deep learning model, Grandio, Riveiro, Soilán, et al. extract cubes with a fixed edge length of 10 m from larger scene [44]. The work of Corongiu et al. extract a cylindrical region (radius=2 m) of interest around candidate points. These cylindrical regions are then further processed to create a semantic segmentation [38] of the scene.

Using information from the scanning geometry and the time-stamp metadata of each point it is possible to extract consecutive cross sections of the railway bed area [68]. These so-called scan lines are then further processed to extract the track locations.

G. FEATURE EXTRACTION

Point clouds offer a rich source of data from which a plethora of features can be derived. Geometrically, one can extract attributes such as normal vectors and curvature. From a statistical perspective, features like local density and variance are valuable. In terms of shape, roughness and linearity provide insights into the structure of the data. Topologically, connectivity sheds light on the relationships between data points. Additionally, when colour information is available, RGB values can be harnessed. These extracted features, encompassing geometric, statistical, shape, topological, and colour attributes, serve as foundational elements for subsequent modelling endeavours.

Geng et al. provide a comparison of several feature extraction methods applied to a point cloud scene of a Chinese high-speed railway collected using an airborne laser scanner [42]. The work of Jung et al. extract line segments per voxel [48]. These line segments are then classified using a multi-range Conditional Random Field (CRF) classifier.

H. OTHERS

The majority of the works use laser scanning techniques to capture a point cloud. An alternative approach is to use photogrammetry techniques to create a point cloud based on image data. This is done in the work of Sahebdivani et al. which use a commercial drone to capture images from the area of interest. These images are then processed to create a point cloud [63]. The use of structured light is another approach to create point clouds, this is done by Cui et al. in their work to automatically inspect railway fasteners [40].

One pre-processing step which is often lacking from literature is the processing of the raw point cloud data. Often laser scanners will produce a stream of frames. These frames are then combined to create a larger point cloud scene. During this processing step, the points are also mapped from their sensor's local reference frame to a global coordinate reference system. To do so, an accurate Global Navigation Satellite System (GNSS) is required. The reception and accuracy of GNSS is not always consistent, therefore GNSS data is often augmented with gyroscope, heading and odometer data. The work of Xu et al. sheds some light on this matter [67]. During the processing of raw frames into larger scenes, also duplicate measurements are excluded. For instance, when the measurement train is standing still, data is still being collected. This will contain a lot of redundant data, which is removed during post-processing.

I. SUMMARY OF PRE-PROCESSING TECHNIQUES

The pre-processing techniques described above are tied closely to the purpose and each of them has its advantages and challenges. The choice of the techniques is mostly dependent

TABLE 3. Comparison of pre-processing techniques for point clouds and their use in the context of railway infrastructure.

Technique	Challenges	Advantages	Reference
Cropping	• Determining optimal boundaries. • Potential loss of important data.	• Reduces data size for faster processing. • Focuses on regions of interest.	[29], [36], [43], [34], [71]
Partitioning	• Deciding optimal partition size. • Handling boundary data between partitions.	• Manages large datasets by breaking them into manageable chunks.	[28], [29], [36], [45]
Normalisation	• Determining appropriate scale. • Potential loss of original data characteristics.	• Standardises data range. • Enhances compatibility with algorithms requiring normalised data.	[28], [38], [61]
Projection	• Loss of 3D information. • Deciding optimal projection plane.	• Converts 3D data to 2D for easier visualisation and processing.	[38], [75]
Data Structures	• Complexity in implementation. • Overhead in memory and computation.	• Efficient data access and manipulation. • Supports advanced algorithms and operations.	[27], [44], [45], [48], [69]
Sampling	• Potential loss of detail. • Choosing appropriate sampling method and density.	• Reduces data size. • Speeds up processing.	[38], [40], [44], [68]
Feature Extraction	• Deciding relevant features. • Complexity in feature computation.	• Reduces data dimensionality. • Enhances data characteristics for specific tasks.	[48], [42]

on the context and the data. In Table 3, we provide a concise summary and comparison of these techniques. We have also included their use in the context of railway infrastructure as a result of our literature study. From the table, it is evident that these pre-processing techniques are not mutually exclusive. Instead, several techniques are often employed to maximise their collective benefit.

V. MODELING TECHNIQUES

In this section, we compile a glossary of methods, algorithms, and techniques for modelling point clouds, designed for purposes like object classification, segmentation, and object detection. We categorise and describe these methods found in the literature, focusing on their strengths and limitations. The point cloud modelling methods are broadly divided into two categories: structure-based methods and machine learning-based methods. We describe each of these and their sub-categorisation. Two aspects are linked to modelling. One is the performance metric, while the other is the type of railway infrastructure being modelled. We start this section by providing information on these two essential aspects.

TABLE 4. Break down of the literature based on railway component.

Rail Component	Reference
Rail Track detection and classification	[26], [40]
Cable detection and classification	[26], [29], [38], [47]
Switch and crossing detection	[31], [32], [34]

A. RAIL INFRASTRUCTURE

An essential aspect to consider in the railway environment is the modelling goal concerning railway infrastructure. While several researchers have focused on specific components of the infrastructure, the complete railway infrastructure is often overlooked. In Table 4 we have summarised the most commonly studied infrastructure components along with the corresponding research references.

It is important to acknowledge that certain aspects of the railway infrastructure, such as foreign objects, bridges, and tunnel deformation, have not been included in this paper due to the set exclusion criteria. Nevertheless, these areas have been gaining interest, particularly in the context of predictive maintenance and the expansion of high-speed rail networks in China (e.g., [37]). As the railway industry continues to evolve, exploring these aspects becomes increasingly crucial for comprehensive railway infrastructure modelling and analysis.

B. PERFORMANCE METRICS

To evaluate the performance of modelling techniques various metrics can be used. In the following, We define the most popular metrics used in the context of point clouds.

1) ACCURACY, PRECISION, RECALL, F_1-SCORE

These are commonly used metrics for evaluating classification accuracy. For the sake of completeness they are defined below:

- **Accuracy** measures the overall correctness of a model's predictions by calculating the ratio of correctly predicted instances to the total number of instances (Equation 1). It provides a general assessment of how well the model performs across all classes. The formula for accuracy is:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where TP stands for true positive, TN is the true negative, FP is false positive, and FN is the false negative.

- **Precision** focuses on the proportion of correctly predicted positive instances out of all instances predicted as positive (Equation 2). It provides insight into the model's ability to avoid false positives (instances predicted as positive but are actually negative). The formula for precision is:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

- **Recall**, also known as sensitivity or true positive rate, measures the proportion of correctly predicted positive instances out of all actual positive instances (Equation 3). It indicates the model's ability to identify all positive instances and avoid false negatives (instances predicted as negative but are actually positive). The formula for recall is:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

- The $\mathbf{F_1}$**-score** is a harmonic mean of precision and recall (Equation 4). It provides a balanced measure that takes into account both precision and recall. The F_1-score is useful when one want to consider both false positives and false negatives equally. The formula for the F_1-score is:

$$F_1\text{-}score = 2\frac{Precision \cdot Recall}{Precision + Recall} \tag{4}$$

In the case of Boolean data, the F_1 score is also sometimes referred to as the Sørensen-Dice coefficient.

2) ROOT MEAN SQUARE ERROR (RMSE)

It is the standard deviation of prediction error (Equation 5). It is often used for regression problems. The formula to compute RMSE is:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Actual_i - Predicted_i)^2}{N}} \tag{5}$$

3) MEAN INTERSECTION OVER UNION

Mean Intersection over Union (mean IoU) is a metric commonly used in evaluating the performance of semantic segmentation models. It measures the overlap between the predicted segmentation and the ground truth segmentation.

The Intersection over Union (IoU), also known as Jaccard Index, for a single class is calculated by dividing the size of the intersection of pixels between the predicted and ground truth masks by the size of the union of those pixels (Equation 6). It provides a measure of how well the model accurately captures the boundaries and regions of the objects of interest.

The mean IoU is then computed by averaging the IoU values across all classes or categories. It provides an overall assessment of the segmentation model's performance, taking into account the accuracy of segmenting multiple classes simultaneously.

The formula for calculating IoU is:

$$IoU = \frac{|Predicted\ mask \cap ground\ truth\ mask|}{|Predicted\ mask \cup ground\ truth\ mask|} \tag{6}$$

where $\cap$ is the intersection, $\cup$ is the union and $|\cdot|$ is the cardinality. The mean IoU is computed by taking the average IoU across all classes or categories (Equation 7). Here, N is the total number of classes:

$$\text{Mean IoU} = \frac{1}{N}\sum_{i=1}^{N} \text{IoU class } i \tag{7}$$

Mean IoU values range from 0 to 1, with 1 indicating a perfect overlap between the predicted and ground truth masks, and 0 indicating no overlap at all. Higher mean IoU values indicate better segmentation performance.

C. STRUCTURE-BASED METHODS

Structure-based methods exploit or enforce structure to the point cloud scenes. These methods utilise the geometric and topological properties of point clouds and often rely on mathematical models to extract meaningful information from the point cloud data.

These methods leverage geometric and topological properties, enabling them to represent the underlying 3D structure and surfaces accurately. Moreover, these methods are frequently characterised by well-defined mathematical models. These models not only enhance our understanding of the underlying processes but also ensure precision during implementation.

Besides their advantages the structure-based methods have limitations too. The methods could struggle to model surfaces that are complex since the underlying principles rely on basic geometric primitives. Additionally, they can handle noise to a certain level but remain sensitive to a high noise level and outliers that can impact their performance and limit their usability. Moreover, structure-based methods could be computationally intense and they do not profit from higher point densities. Their lack of adaptability is another drawback, as they are often tailored to specific application domains and may not generalise well to diverse datasets.

These methods can be further categorised (see e.g. [8], [78]). The following subsections describe these sub-categories and their use in the context of the railway environment.

1) EDGE-BASED METHODS

Edge-based methods usually have two main stages: (i) detecting edges to outline borders of different regions followed by (ii) the grouping of points inside boundaries to generate the final segments. Edges are defined by points where changes in the local surface properties exceed a given threshold. Local surface properties are for instance normals, gradients, principal curvatures or higher-order derivatives. Edge-based methods are generally fast but may produce inaccurate results in case of noise and uneven density of point clouds. When disconnected edges are detected, a filling or interpretation procedure is applied to identify closed segments. Ariyachandra and Brilakis [28] used this method to detect pole-like objects. Their approach was based on line detection and point clustering.

2) REGION-GROWING METHODS

Region-growing methods start from one or more seed points that possess specific characteristics and then expand to neighbouring points with similar characteristics. These characteristics are for example surface orientation, curvature, etc. Bottom-up approaches start from some seed points and grow the segments on the basis of given similarity criteria. Bottom-up region-growing algorithms include two steps: identification of the seed points and adding points to them based on predefined criteria. Top-down approaches start by assigning all points to one segment and then subdivide the segment into smaller ones guided by certain thresholds. Region-growing methods are robust to noise (see e.g. [79]), but they are sensitive to (i) the location of initial seed regions and (ii) inaccurate estimations of the normals and curvatures of points near region boundaries. For an explicit description of the region-growing algorithm in the context of railways, the interested reader is referred to Cserép et al. [39, Algorithm 1 and Algorithm 2]. Other examples of region-growing algorithms can be found in the work of Arastounia [25], Chbeir et al. [34], Zhang et al. [24], Lu et al. [23], and Zou et al. [71].

3) MODEL FITTING METHODS

Model fitting methods are based on the observation that a lot of objects are built-up out of geometric primitives like planes, cylinders and spheres. Primitive shapes are fitted onto the point cloud and the points that comply with the mathematical representation of the primitive shape are labelled as one segment. Widely employed algorithms for model fitting are Hough Transform (HT), Random Sample Consensus (RANSAC) and fast point feature histograms (FPFH). Note that HT and FPFH are used to generate features that are utilised as an input for the model fitting methods such as RANSAC (see e.g. [80]). Model fitting methods are fast and robust with outliers. However, they fall short when dealing with complex shapes or fully automated implementations. Moreover, they have problems when dealing with different scales of input point clouds. References that utilise these techniques in the context of railways are [26], [29], [30], [31], [60], [61]. A comparison of these methods is presented in [39].

4) GRAPH-BASED METHODS

Graph-based methods view point clouds as graphs. In the simplest model, the vertices in the graph correspond to points in the data and the edges represent certain pairs of neighbouring points [81]. An alternative approach is to first aggregate points into coherent patches, these patches are then considered as the vertices of the graph [82]. Other techniques first voxelise the cloud with for example an octree or supervoxel method and construct a graph out of the voxelised point cloud. Graph-based methods are able to segment complex scenes in point cloud data with noise or uneven density with good results for example by finding the minimum-cut of the graph. However, these methods usually can not run in real-time and some of them may need an offline training step.

Although the technique is applied in other contexts it is not applied in the context of railway environments.

5) HYBRID METHODS

Multiple different methods are combined to exploit the best parts of the methods. Most of the reviewed papers fall under this category. Examples include [26], [28], [29], [34], [51]. Zhang et al. has used several algorithms to extract power lines. They used the spatial structure of the power line for initial segmentation followed by a region-growing method. They applied PCA on the results of the region-growing algorithm and as a final step, they used least square fitting algorithm to model power lines [24]. Zou et al. has used a combination of k-mean clustering and region-growing algorithm to extract railway tracks from point cloud data. Their focus was on extracting railway tracks with complex topology like bends and turnouts [71].

6) SUMMARY OF THE STRUCTURE-BASED METHODS

All structure-based methods have their strong and weak points. In Table 5, we have tabulated the advantages and challenges associated with each of these methods. We have also included the use of these methods in the context of the railway environment based on our literature search.

D. MACHINE LEARNING-BASED TECHNIQUES

Machine learning algorithms owe their success to their ability to learn from data. The popularity and widespread adoption of these algorithms can be largely attributed to the vast availability of data. Unlike structural methods, machine learning approaches are inherently adaptive, autonomously discovering patterns in the data. The performance of machine learning methods heavily relies on the quantity and quality of the training data.

Various machine learning techniques have been developed specifically tailored to the unique structure of point clouds. As mentioned before, working with point cloud data presents its challenges, and the degree of success achieved by these algorithms is often limited. Many of these algorithms struggle to generalise well to different domains. Unlike computer vision problems, where large pre-trained models are readily available for transfer learning, point cloud tasks lack such widespread pre-trained resources. Besides its limitation, the use of machine learning-based techniques is trending, which will become apparent in the following sections.

The machine learning-based techniques can be broadly categorised as traditional machine learning-based techniques and deep learning techniques.

TABLE 5. Comparison of structure-based methods for point clouds.

Method	Advantages	Challenges	References
Edge-based	• Effective in detecting sharp features and boundaries • Useful for feature extraction and edge detection	• Sensitive to noise and outliers • Limited in handling non-sharp transitions	[28]
Region Growing	• Generates coherent and connected regions • Robust to noise and capable of handling irregular shapes	• Sensitivity to seed point selection • May struggle with complex structures	[24], [25], [34], [71]
Model Fitting	• Provides accurate geometric representations • Useful for shape recognition and surface reconstruction	• Prone to model selection bias • Computational complexity for complex models	[26], [29], [31]
Graph-based	• Captures geometric relationships effectively • Suitable for segmentation, classification, and features	• Complexity in designing and training GCNs • Requires graph construction and processing overhead	–
Hybrid	• Combine strengths of multiple methods to leverage complementary information	• Memory and computational overhead could be higher • Integration and hyperparameter tuning could be a challenge	[24], [26], [28], [29], [34], [51], [71]

1) TRADITIONAL MACHINE LEARNING

Traditional machine learning methods are employed as an evolution towards methods that learn from the data and have better generalisability. The main difference with deep learning methods is that deep learning methods learn the features themselves. This is generally faster and more efficient since the model can derive more complex features and can distinguish the most informative features.

The literature on traditional machine learning in the context of the railway environment is currently limited.

Sturari et al. have presented a traditional machine learning approach. The primary focus was on data collection methodology, and to provide proof of concept for applicability of machine learning to the collected dataset. They have compared four traditional machine learning methods, namely decision tree, support vector machine, k-nearest neighbour and random forest with a convolution neural network. Notice that the learning task differed since the authors were concerned with the classification problem instead of segmentation or object detection [22].

The approach used by Uggla and Horemuz is interesting since they combined synthetic and real-world data to create synthetic railway scenes. They compared the performance of the deep learning-based approach on real scenes and the scenes synthetically generated from a point cloud. They concluded that the synthetic data could be helpful in generalising performance since more data can be generated easily instead of going through a lengthy data collection process [65].

2) DEEP LEARNING BASED METHODS

The problem of point cloud segmentation and classification looks similar to the ones in computer vision. However, the segmentation and classification are much more challenging for point clouds due to the absence of the grid structure. In computer vision, the images are represented using a structured grid of pixels that allows the application of a convolution neural network (CNN) to extract features. This structure is not present in point clouds. Thus, CNN cannot be directly applied.

The point clouds are irregular and sparse since the density and distribution of points can vary significantly between different scenes and objects.

Deep learning on point clouds requires a distinct approach, and it is currently attracting increasing interest from researchers, particularly in the context of self-driving cars [6]. For an empirical comparison of various deep learning approaches, the interested reader is referred to [83]. It is worth mentioning that Guo et al. conducted an empirical comparison of existing deep learning-based approaches for 3D point clouds using various benchmark datasets. However, it is important to note that none of these benchmark datasets specifically include the railway environment.

The deep learning-based methods can be broadly classified into three categories based on how they handle the point clouds. The categories, indirect methods, direct methods and hybrid methods are addressed in the subsequent sections in context of the railway environment.

a: INDIRECT METHODS

Indirect methods rely on the volumetric representation of point clouds before applying deep learning techniques. Due to this volumetric representation, these methods are also called volumetric or grid-based methods. The most often used volumetric representations are voxel clouds, octree, and projections. The main idea behind the volumetric representation is to introduce a structure similar to images to implement the neural network architecture similar to computer vision. Furthermore, the volumetric representation facilitates the use of 3D convolutions and learning global context. These methods are also limited since creating volumetric representation introduces an extra computational overhead. In the case of voxelisation, the performance depends on voxel size since large voxel size may lead to loss of fine-grained information.

Indirect methods are often used in conjunction with the convolutional neural network. The fundamental difference among these methods is the underlying volumetric representation. Example includes octrees (OctNet [84], SPH3D-GCN [85], O-CNN [86]), voxel clouds [87] and *kd*-trees (Kd-network [88]).

In the context of railway infrastructure, only a few papers using three-dimensional CNNs were found during our literature search. Lin et al. has used CNN to identify the context of each single frame point cloud [54]. On the other hand, Corongiu, Masiero, and Tucci have used CNN with modified Fisher vectors for object classification and extraction of geometrical features with support vector machines [38]. Yu et al. has used voxelised data as input for the deep learning algorithm. Their architecture is based on 3D CNN with a shared MLP and max pooling [69].

Another way to create a volumetric representation is to use projections. The projections convert the point clouds into 2D images, thus enabling 2D convolution. In the context of railway infrastructure, Manier et al. has used a projective descriptor together with neighbourhood selection for point cloud classification. They focused on computational speed gains though convergence is not guaranteed for challenging datasets [58].

b: DIRECT METHODS

Another approach to handle point cloud data is to use raw point clouds directly. Since these methods directly utilise point clouds the computational- and memory costs for intermediate representation, computation, and storage are saved. Also, the direct use of point clouds facilitates non-uniform density and spatial distribution. Moreover, these methods can capture fine-grained point-level information, which is not the case for indirect methods.

These methods have certain limitations. Due to the large size of point clouds, these approaches are more computationally involved. Another limitation is the lack of context information since the neural network works directly with the raw point clouds. However, they handle irregularity better.

PointNet is one of the pioneering methods for directly processing point clouds [20] (see also [83] for an empirical comparison). The method has a shortcoming since it does not consider the local relationships within the point cloud. It uses a shared multi-layer perceptron (MLP) and a symmetric function to aggregate information from individual points, resulting in a global feature vector representing the entire point cloud. However, creating this feature vector is solely based on individual points resulting in a loss of contextual or global information. PointNet++ is an evolved version of PointNet since it also considers the hierarchical relationship among different points reducing the computational requirement and increasing its accuracy.

PointNet++ is based on shared MLP and max pooling to introduce the concept of local neighbourhood thus capturing both local and global information. It is widely adopted in the context of railway scene classification/segmentation, examples include [44] that have used it together with random subsampling while Dibari et al. [41] have used it in conjunction with transfer learning for semantic segmentation.

c: HYBRID METHODS

Hybrid methods combine direct and indirect approaches, incorporating raw point cloud data and some intermediate representation. These methods often leverage the benefits of both paradigms to achieve improved performance and efficiency. The most often used terms in this respect are point fusion and projection-based methods. The projection-based methods are also classified as an indirect method depending on whether the projection is used to introduce structure or in conjunction with raw point clouds. Hybrid methods need to effectively combine information from direct and indirect representations, requiring careful design to ensure the fusion process does not introduce artifacts or redundancies. Some hybrid methods may require storing both raw point clouds and intermediate representations, leading to higher memory usage.

Liu et al. have designed a lightweight neural network with an attention mechanism. The attention mechanism was designed to concentrate on important features ignoring the unimportant ones mimicking human cognition [55].

The approach adopted by Chen et al. can be considered as a hybrid approach. They aimed to consider the point cloud data as sequential data mimicking the scan line view of railway infrastructure. They have used a point partitioning algorithm to determine the region of interest for extracting features. The points in the region of interest are then used to create a neural network-based architecture. The author designed a multi-layer neural network architecture with PointNet as one of the layers. To capture sequential information, they used a form of recurrent neural network [36].

d: SUMMARY INDIRECT, DIRECT, AND HYBRID METHODS

The three classes have their own advantages and disadvantages and their use is tied to the context. In Table 6 we have summarised these methods with respect to their advantages and challenges. We have also included references to the use of these methods in the context of railway environment.

e: FUSION WITH IMAGES AND DEEP LEARNING

Point clouds are used mostly in the setting when the collection of image data is not always feasible. In some settings image data is also available. Thus it is intuitive to apply computer vision techniques to this dataset. In the scenario when both image and point cloud data are available, one can fuse them together to reap benefits from the structure of both data.

A notable work in this direction is presented by [66]. They combined image data and point cloud data to develop an attention mechanism-based algorithm where the

TABLE 6. Summary of direct, indirect, and hybrid methods and reference to their use in the context of railway environment.

Method	Challenges	Advantages	References
Indirect	• Volumetric representation overhead • Spatial resolution limitation	• Regular grid processing • Efficient convolution on regular grids	[38], [54], [58], [69]
Direct	• Irregular data representation • High computational cost	• Efficient data representation • Flexibility in handling varying data characteristics such as density	[36], [41], [44]
Hybrid	• Fusion complexity • Increase memory consumption • Discretisation artefacts	• Global context • Improved performance • Flexibility in balancing efficiency and accuracy	[55]†

†No exact reference found in the context of railway infrastructure modelling.

transformation is performed by spherical projection. Similarly, Wolf, Richter, and Döllner have converted point clouds into images. They used the well-known YOLO [76] framework which can be used for image segmentation and classification.

f: FUSION WITH OTHER DATA AND DEEP LEARNING
Generally, the data collection is not limited to only one sensor instead multiple sensors are used simultaneously to collect various types of data. These data may include images, GPS data, inertial measurement unit, and mapping information. Fusing these datasets into a working methodology could help improve the performance of segmentation tasks. Mahtani et al. have used GNSS data, IMU data and point clouds to develop a semantic segmentation algorithm for railway infrastructure. For the point clouds they have used KP-FCNN [89] which is based on point convolution that does not require any intermediate representation.

E. OTHER TECHNIQUES
1) SEMANTIC AND ONTOLOGY BASED METHODS
These methods incorporate semantic and ontological information of point cloud data to enhance the understanding and classification of 3D point cloud data. These methods aim to assign semantic labels to individual points or segments of the point cloud, indicating the object category or class they belong to. Karmacharya, Boochs, and Tietz have used semantics for object annotation. They combined numerical techniques (structure-based methods (see Section V-C)) with expert domain knowledge to develop inference rules. The rules are used to annotate objects of interest from the point clouds [50]. On the other hand, [33] have used a three-stage approach to use semantic information for object detection and classification.

2) SOFTWARE-BASED APPROACH
Detection of rails using MLS, building detection using polygons. The work is based on existing software tool TerraScan [52].

F. COMMERCIAL SOFTWARE
There has been an uptake of point cloud based information extraction for the railway domain by commercial software vendors. This is a good indicator of the domain becoming more mature. A small, and by no means complete, desk study has been conducted to evaluate the current state and possibilities provided by these commercial software vendors. A total of six software vendors have been compared, and a summary is provided in Table 7.

Half of the products provide a digital terrain model (DTM) of the ballast. Terrasolid states that is capable of detecting missing ballast, though it remains unclear what exactly is meant with this. Another oddity is claim of Leica about measuring ballast volume. This is not possible based on point cloud data alone. The majority of the products are only capable of extracting the wires from the point cloud. Only two products provide detailed information about the wires such as height and stagger. Also, most product provide clearance measurement options.

Products from Leica and TopoDOT provide (semi)-automated object detection algorithms to detect objects such as platforms, poles, and signals. A fully automatic segmentation of the scene is provided by TheCrossProduct and is based on a deep learning approach. Most of the commercial software product also support tunnel deformation measurements. Another unique feature advertised by the Terrasolid product is the risk evaluation of trees adjacent to the track. It evaluates the risk of trees falling on the track or within the clearance space.

VI. DIGITAL TWIN
Digital twins can be defined as a bidirectional effortless data integration between a physical and virtual machine [90]. The terminology is defined extensively and is often confused with the notion of a static digital model, which is a mere virtual representation of the physical world [90]. However, as opposed to static models, digital twins are a dynamic digital representation of a physical object or system. This goes beyond just a mere 3D representation to include real-time data which allows the digital twin to simulate, predicate and optimize its physical counterpart's performance and operation. This of course costs a continuous flow of data from the physical object to maintain the accurate simulation and analysis. Throughout the lifetime of a digital twin, a real-time mirroring is required to reflect any changes or updates, which makes it a powerful tool in understanding, analyzing and improving real-world objects and systems.

TABLE 7. A comparison of features offered by commercial software vendors.

Product	Domicile	Ballast	Track	Wires	Clearance	Objects
Atlas computers - SCC	Ireland	DTM	curvature, vert. grad., gauge, cant, slew, twist	height & stagger	N	N
Leica Geosystems Rail:Factory and ATtrack	Switzerland	volume and DTM	curvature, vert. grad., gauge, cant	extraction only	Y	Y
TheCrossProduct	France	N	curvature, vert. grad., gauge, cant, twist, warp	height & stagger	Y	Y
TopoDOT	United States	DTM	gauge	extraction only	N	Y
Terrasolid	Finland	missing ballast detection	curvature	extraction only	Y	N
SITECO Informatica Rail-SIT	Italy	N	curvature	extraction only	Y	N

The use of digital twins is gaining traction in the context of railway infrastructure maintenance and monitoring. However, the scientific literature is scarce and the focus is on creation of digital model of railways with emphasis on 3D model generation and integration.

In the context of digital twins, 3D model generation refers to the process of creating a 3D representation of the physical objects or environment in a digital twin environment. Incorporating 3D models enables realistic and immersive visualisation of railway assets allowing an assessment of the current state and prediction of the future state.

The earliest work in the included studies is that of Ben Hmida, Cruz,. They have created VRML files with different coloured object classes using an ontological approach. Ariyachandra and Brilakis have focused on generating Industry Foundation Classes (IFC) models. In their first paper, they fitted the detected point clusters into 3D models in IFC format [28]. While the second paper was focused on generating dynamic IFC models of overhead line equipment configurations and merging them with point clusters using the iterative closest point algorithm [29]. Soilán, Sánchez-Rodríguez et al. has reported work on generating IFC files for track models and rail alignment (see also [64]).

The other approaches were based on geometric-based modelling with curve fitting. The studies employing curve-fitting approaches include:

- Parameter estimation using Markov Chain Monte Carlo and curve fitting for rail track modelling [60].
- Rotational correction and projection of 3D points, fitting a pre-defined rail model, and interpolation using Fourier curve fitting [63].
- Piecewise straight line fitting for contact wire and dropper [67].
- Identifying and classifying railway cables (contact, catenary, return current) [37].
- Track model reconstruction using a third-degree polynomial function [49].
- Hybrid overlay technique using point data and polygons [52].
- Use of particle swarm optimisation to reconstruct the track considering the so-called track lining distance as an evaluation index [53].
- Heuristic-based point cloud pre-processing is used to segment railway to generate 3D model based on IFC requirements [46].

In the context of model generation Zhu and Hyyppa have developed a complete building model by fusing facades with roofs, planar detection of buildings, ground model simplification, orthophoto for ground texture, and 3Ds Max for model visualisation.

Despite the scarcity of the literature in the application of digital twin technology in railway infrastructure, a wider adoption is expected to revolutionize railway monitoring and optimization. Digital twin has the potential of enhancing predictive maintenance enabling proactive issue resolution. This technology will not only optimize railway operations but also significantly improve its safety through simulations and hazard identification.

VII. DISCUSSION

In this section we summarise our findings and highlight potential challenges and research directions.

A. PRE-PROCESSING TECHNIQUES

Most of the pre-processing techniques are aimed at reducing the volume of data. Since current models cannot handle large volumes of data due to memory requirements or computational resources. Most post-processing techniques are based on heuristic methods, where parameters such as partition sizes and down-sampling ratios are empirically defined. A possible approach to address this issue is to explore new and innovative models which can handle large volumes of data by themselves. These models could learn to focus on the relevant parts of the input, and might be able to learn sparse structures which fit the underlying data. These approaches could be encoded as a dedicated head of a machine learning model which can be easily added to existing models. An alternative would be to use the raw frame-by-frame data which avoids the issue of partitioning the data into tractable pieces altogether. Furthermore, the format which point cloud data is stored can be enhanced. One such enhancement is to explore formats where the level of detail can be automatically encoded.

TABLE 8. A summary of modelling techniques with related references.

Modelling Technique	References
Structure based-methods	
Edge Based Methods	[28]
Region Growing methods	[25], [34], [39]
Model Fitting Methods	[26], [29], [31], [39]
Graph Based Methods	-
Hybrid Methods	[26], [28], [29], [34], [51]
Machine learning-based methods	
Traditional Machine Learning	[22], [65]
Deep Learning Based Methods: Indirect method	[38], [54], [58], [69]
Deep Learning Based Methods: Direct method	[36], [41], [44]
Deep Learning Based Methods: Hybrid method	[55]
Other techniques	
Semantic and Ontology based	[50]
RANSAC based Methods	[60], [61]
DBSCAN	[23]

This would aid the construction of models which work from coarse detection towards fine detection. For instance, detecting locations of catenary arches would require a low level of detail, whereas locating the exact location of insulators within the scene would require a high level of detail.

B. MODELLING TECHNIQUES

Machine learning-based methods have gained significant importance in railway point classification as shown in Table 8, with seven publications using it for point cloud classification or segmentation. Though the models involved usually require large training datasets of point cloud data to be able to learn patterns, features and relationships, the efforts are worthwhile in terms of performance and allow for transferability to other problems (model reuse or transfer learning). Convolutional and graph neural nets dominate the machine learning models used in the various problems related to railway point clouds, such as object detection or rail track detection. Machine learning approaches are also powerful in the sense of their generalisability to different railway environments and condition on top of the availability of pre-trained models ready for reuse under minor fine-tuning. The drawback of such approaches remains their limited explainability, especially for transformer-based and recurrent models.

The runner-up technique is model fitting, which involves fitting geometric models to the point cloud data, enabling the identification of rail tracks, switches, poles, or other relevant structures. Such techniques have been used in four of the surveyed works. The technique is powerful in the sense that it allows to explicitly model and detect specific railway objects and structures, as well as its robustness against outliers. Its main drawback remains however the labour-intensive effort to initialise and set the parameters for the models with due care to the calibration to optimise the results. Such calibrations efforts should of course factor in the various railway environments and conditions.

C. RESEARCH ROADMAP

We have condensed the information and insights from this literature review into a research roadmap. This roadmap indicates several pointers for future research. First the need for a benchmark dataset is put forward, together with the requirements of such a dataset. After this we focus on the advancements within the area of machine learning which are needed to efficiently and effectively deal with large volumes of point cloud data. Lastly we provide indicators within the area of usability, such as visualisation and interoperability.

1) BENCHMARK DATASET

We like to emphasise the need for an open benchmark dataset to rank various techniques proposed in literature. The number of publications in the context of point cloud, machine learning and railway infrastructure is increasing. However, one cannot objectively compare these techniques due to a lack of a common benchmark dataset. We again emphasise developing a benchmark dataset set comparable to the KITTI dataset [92] available for self-driving cars. The dataset published by Ton [93] is a positive step in this direction. We propose the creation of a public benchmark dataset as a joint collaborative effort between academics and industry. Goal will be to create dataset with a large variety of railway scenes. Within a single country the railway components can vary significantly in appearance. Ideally the dataset should have a large inter-country and intra-country variation. The dataset should also support multiple machine learning tasks such as segmentation and object detection. Regarding object detecting, the dataset should also encode the directionallity of the object. Furthermore a hierarchical labelling scheme is recommended, the dataset should facilitate detecting larger scale object such as catenary arches, but also individual components such as insulators. From previous experience it was noted that each labelled object should also have a unique identifier, this aids the automated iteration of objects.

2) MACHINE LEARNING ADVANCEMENTS

A focus area of active research could be the creation of lightweight machine learning models which are able to consume raw point clouds directly. Current machine learning models are not capable of this yet.

a: WEAKLY SUPERVISED LEARNING

Machine learning-based techniques and structure-based methods have shown promising performance. However, the cost of data labelling hampers large-scale application of machine learning-based techniques. Data annotation is a tedious, time-consuming, and error-prone task. Applying various machine learning techniques that work with partially labelled data is advisable. Our research group has had some success in this direction (see [94]), hindered mainly by the data labelling process. We have worked with an active learning approach; the results will be published elsewhere.

We emphasise that the scale of our experiments are limited, and there is a need for larger-scale experimentation to quantify the strength and feasibility of applying machine learning techniques for partially labelled data.

Another promising direction is self supervised learning (see e.g. [95]). Often, there are large volumes of unlabelled data available and very few labelled examples. Goal would be to leverage this huge volume of unlabelled data to encode some ‘common sense‘ into a model. Knowledge of this model can then be transferred when training a new model based on labelled data. This technique has been applied to point clouds for other domains but to the best of our knowledge has not been applied to railway datasets yet.

b: MULTI-MODAL AND ENSEMBLE MACHINE LEARNING

The railway environment can be captured using different data modalities that can be combined for an optimal performance. Fusion of image data with point clouds have shown promising results (see e.g. [57], [66], [75]). Different data modalities can be combined systematically to develop a multimodal-model. However, this approach has its own challenges [96]. This approach is already used for point cloud data (see [97]) however it is not applied in the context of railway environment.

Another challenge in railway data is relative sizes of objects that hampers development of a single model that captures all variation. These could also be handled by multimodal-model approach where variation is used as a modality for modelling. Another way to handle variation is develop separate models based on object sizes and ensemble them.

c: HYBRID APPROACH

Structural methods exploit topological structure, and the railway environment has fixed topology to some extent. A combination of machine learning and a structural approach could lead to promising results in the context of railway infrastructure. Based on our experiments with various structural and machine learning-based methods, for objects with defined shapes like track lines and catenary cables, structural approaches could be more fruitful for complex objects such as an insulator or signals machine learning-based approaches could be promising. Thus, a combination of both has the potential to optimise accuracy for segmentation and object detection tasks.

d: EXPLAINABLE AI

The other aspect is the application of explainability. Current machine learning models for point clouds are mostly black boxes. In our literature search we cannot find a reference where explainablity is applied to point cloud data irrespective of application domain. Since the point cloud data is distinct in its working, there is a need to develop explainable AI techniques tailored towards point cloud data. As a first step one can evaluate the applicability of the so-called model agnostic approaches (see e.g. [98]) for point cloud datasets.

3) USABILITY

The final research direction is the usability of the results obtained from the data. Questions arise, such as: How to visualise and interact with the results, how to ensure interoperability of the results?

The world around us is three-dimensional, but still the most common way to visualise data captured from this three-dimensional world is on a two-dimensional screen. Recent technological advancements within the area of head-mounted displays (HMD) have enabled very interesting opportunities to visualise and interact with 3D data (see e.g. [99]). This technology can also be used to interactively label the data and to visualise the results from the machine learning models.

As the information which is extracted from point cloud data will likely be used in a larger context, the interoperability of this information is vital. Without proper interoperability, there is a risk of isolated digital environments being formed [100]. To overcome this risk, techniques such as linked data [101] can be explored. These techniques can leverage information derived from point clouds into broader contexts, such as asset monitoring [102].

VIII. CONCLUSION

This paper has reviewed the literature on using point clouds in the context of railway infrastructure. We have divided the literature into pre-processing, modelling, and digital twin creation. We have described different techniques describing their strengths and weaknesses. We have focused on literature concerning railway infrastructure and point clouds, thus excluding literature studying the presence of foreign objects in railway infrastructure that could be an exciting topic concerning predictive maintenance.

The current trend for modelling is focused on machine learning-based techniques, particularly deep learning-based techniques. However, contrary to the usual practice in AI research, data and implementation code is not published for most research, hindering reproducibility and cross-comparison. We emphasise a need towards open publication of data and implementation for scientific research, enabling breakthroughs in this area of research.

Besides the typical challenges associated with point cloud data, railway data have additional challenges, such as variation in object types and sizes, vast sizes of data, and the critical nature of infrastructure hindering the open publication of point cloud data.

As a future research direction, we propose to focus on hybrid methods to reap benefits from the strengths of structure-based and machine learning-based techniques. Also, a focus on developing large pre-trained models for point clouds, in general, will enable transfer learning that reduces the training efforts. Similarly, machine learning techniques focused on partially labelled data can improve the state of the art. From the perspective of digital twins, in the context of railway infrastructure and point cloud, there are still many open research opportunities.

To conclude, the current state of the art is varied in terms of techniques and technologies and could be further strengthened with the use of hybrid methods, multimodal-model and ensemble learning approaches, partially labelled data approaches, and the creation of digital twin to reap full-scale benefits of predictive maintenance enabled and digitalised railways.

REFERENCES

[1] I. Pagand, "Fostering the railway sector through the European Green Deal," Eur. Union Agency Railways, Valenciennes, France, Tech. Rep. ERA1234, 2020.

[2] *Shift2Rail Strategic Master Plan*, Eur. Commission, Brussels, Belgium, 2015.

[3] M. Ruete, "ERTMS: First work plan of the European Coordinator," Eur. Commision, Brussels, Belgium, Tech. Rep. 13/2017, 2020.

[4] J. M. Easton, "Digitalization and the transport industry—Are we ready?" in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2022, pp. 3814–3820.

[5] F. Di Stefano, S. Chiappini, A. Gorreja, M. Balestra, and R. Pierdicca, "Mobile 3D scan LiDAR: A literature review," *Geomatics, Natural Hazards Risk*, vol. 12, no. 1, pp. 2387–2429, Jan. 2021.

[6] Y. Li, L. Ma, Z. Zhong, F. Liu, M. A. Chapman, D. Cao, and J. Li, "Deep learning for LiDAR point clouds in autonomous driving: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 8, pp. 3412–3432, Aug. 2021.

[7] K. Mirzaei, M. Arashpour, E. Asadi, H. Masoumi, Y. Bai, and A. Behnood, "3D point cloud data processing with machine learning for construction and infrastructure applications: A comprehensive review," *Adv. Eng. Informat.*, vol. 51, Jan. 2022, Art. no. 101501.

[8] A. Nguyen and B. Le, "3D point cloud segmentation: A survey," in *Proc. 6th IEEE Conf. Robot., Autom. Mechatronics (RAM)*, Nov. 2013, pp. 225–230.

[9] L. Yong, G. Tong, J. Yang, L. Zhang, H. Peng, and H. Gao, "3D point cloud scene data acquisition and its key technologies for scene understanding," *Laser Optoelectron. Prog.*, vol. 56, no. 4, 2019, Art. no. 040002.

[10] M. Corongiu, G. Tucci, E. Santoro, and O. Kourounioti, "Data integration of different domains in geo-information management: A railway infrastructure case study," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 4, pp. 121–127, Sep. 2018.

[11] Z. Dong, F. Liang, B. Yang, Y. Xu, Y. Zang, J. Li, Y. Wang, W. Dai, H. Fan, J. Hyyppä, and U. Stilla, "Registration of large-scale terrestrial laser scanner point clouds: A review and benchmark," *ISPRS J. Photogramm. Remote Sens.*, vol. 163, pp. 327–342, May 2020.

[12] M. J. Pappaterra, F. Flammini, V. Vittorini, and N. Bešinović, "A systematic review of artificial intelligence public datasets for railway applications," *Infrastructures*, vol. 6, no. 10, p. 136, Sep. 2021.

[13] A. Klimkowska, S. Cavazzi, R. Leach, and S. Grebby, "Detailed three-dimensional building façade reconstruction: A review on applications, data and technologies," *Remote Sens.*, vol. 14, no. 11, p. 2579, May 2022.

[14] Veritas Health Innovation. *Covidence Systematic Review Software*. Accessed: Feb. 2, 2023. [Online]. Available: https://www.covidence.org

[15] Elsevier. *Scopus*. Accessed: Feb. 2, 2023. [Online]. Available: https://www.scopus.com

[16] Clarivate. *Document Search—Web of Science Core Collection*. Accessed: Feb. 2, 2023. [Online]. Available: https://www.webofscience.com

[17] Association for Computing Machinery. *ACM Digital Library*. Accessed: Feb. 2, 2023. [Online]. Available: https://dl.acm.org

[18] Schloss Dagstuhl—Leibniz-Zentrum für Informatik GmbH. *DBLP: Computer Science Bibliography*. Accessed: Feb. 2, 2023. [Online]. Available: https://dblp.org

[19] Institute of Electrical and Electronics Engineers. *IEEE Xplore*. Accessed: Feb. 2, 2023. [Online]. Available: https://ieeexplore.ieee.org

[20] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 77–85.

[21] B. Ton, F. Ahmed, and J. Linssen, "Semantic segmentation of terrestrial laser scans of railway catenary arches: A use case perspective," *Sensors*, vol. 23, no. 1, p. 222, Dec. 2022.

[22] M. Sturari, M. Paolanti, E. Frontoni, A. Mancini, and P. Zingaretti, "Robotic platform for deep change detection for rail safety and security," in *Proc. Eur. Conf. Mobile Robots (ECMR)*, Sep. 2017, pp. 1–6.

[23] Z. Lu, Z. Zhu, J. Bi, K. Xiong, J. Wang, C. Lu, Z. Bao, and W. Yan, "Bolt 3D point cloud segmentation and measurement based on DBSCAN clustering," in *Proc. China Autom. Congr. (CAC)*, Beijing, China, Oct. 2021, pp. 420–425.

[24] S. Zhang, C. Wang, Z. Yang, Y. Chen, and J. Li, "Automatic railway power line extraction using mobile laser scanning data," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 43, pp. 615–619, Jun. 2016.

[25] M. Arastounia, "An enhanced algorithm for concurrent recognition of rail tracks and power cables from terrestrial and airborne LiDAR point clouds," *Infrastructures*, vol. 2, no. 2, p. 8, Jun. 2017.

[26] M. Arastounia and S. O. Elberink, "Application of template matching for improving classification of urban railroad point clouds," *Sensors*, vol. 16, no. 12, p. 2112, Dec. 2016.

[27] M. Arastounia, "Automated recognition of railroad infrastructure in rural areas from LiDAR data," *Remote Sens.*, vol. 7, no. 11, pp. 14916–14938, Nov. 2015.

[28] M. R. M. F. Ariyachandra and I. Brilakis, "Detection of railway masts in airborne LiDAR data," *J. Construct. Eng. Manage.*, vol. 146, no. 9, Sep. 2020, Art. no. 04020105.

[29] M. R. M. F. Ariyachandra and I. Brilakis, "Digital twinning of railway overhead line equipment from airborne LiDAR data," in *Proc. 37th Int. Symp. Autom. Robot. Construction (ISARC)*, Oct. 2020, pp. 1270–1277.

[30] R. Beger, C. Gedrange, R. Hecht, and M. Neubert, "Data fusion of extremely high resolution aerial imagery and LiDAR data for automated railroad centre line reconstruction," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 6, pp. 40–51, Dec. 2011.

[31] H. Ben Hmida, C. Cruz, F. Boochs, and C. Nicolle, "From 3D point clouds to semantic objects: An ontology-based detection approach," in *Proc. Int. Conf. Knowl. Eng. Ont. Devel. (IC3K)*, vol. 1, Oct. 2011, pp. 255–260.

[32] H. Ben Hmida, C. Cruz, C. Nicolle, and F. Boochs, "Knowledge-driven method for object qualification in 3D point cloud data," in *Advances in Knowledge-Based and Intelligent Information and Engineering Systems*, vol. 243. Amsterdam, The Netherlands: IOS Press, 2012, pp. 258–267.

[33] H. Q. Truong, H. B. Hmida, F. Boochs, A. Habed, C. Cruz, Y. Voisin, and C. Nicolle, "Automatic detection and classification of objects in point clouds using multi-stage semantics," *Photogrammetrie Fernerkundung Geoinf.*, vol. 2013, no. 3, pp. 221–237, Jun. 2013.

[34] J.-J. Ponciano, C. Prudhomme, B. Tietz, and F. Boochs, "Detection and isolation of switches in point clouds of the German railway network," in *Proc. 11th Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, Nov. 2015, pp. 96–102.

[35] Y.-J. Cheng, W.-G. Qiu, and D.-Y. Duan, "Automatic creation of as-is building information model from single-track railway tunnel point clouds," *Autom. Construct.*, vol. 106, Oct. 2019, Art. no. 102911.

[36] L. Chen, C. Xu, S. Lin, S. Li, and X. Tu, "A deep learning-based method for overhead contact system component recognition using mobile 2D LiDAR," *Sensors*, vol. 20, no. 8, p. 2224, Apr. 2020.

[37] X. Chen, Z. Chen, G. Liu, K. Chen, L. Wang, W. Xiang, and R. Zhang, "Railway overhead contact system point cloud classification," *Sensors*, vol. 21, no. 15, p. 4961, Jul. 2021.

[38] M. Corongiu, A. Masiero, and G. Tucci, "Classification of railway assets in mobile mapping point clouds," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 43, pp. 219–225, Aug. 2020.

[39] M. Cserép, A. Demján, F. Mayer, B. Tábori, and P. Hudoba, "Effective railroad fragmentation and infrastructure recognition based on dense LiDAR point clouds," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 2, pp. 103–109, May 2022.

[40] H. Cui, J. Li, Q. Hu, and Q. Mao, "Real-time inspection system for ballast railway fasteners based on point cloud deep learning," *IEEE Access*, vol. 8, pp. 61604–61614, 2020.

[41] P. Dibari, M. Nitti, R. Maglietta, G. Castellano, G. Dimauro, and V. Renò, "Semantic segmentation of multimodal point clouds from the railway context," *Proc. SPIE*, vol. 11785, Jun. 2021, Art. no. 117850S.

[42] Y. Geng, Z. Wang, L. Jia, and Y. Qin, "Comparison of LiDAR point cloud features in railway environment," in *Proc. IEEE Int. Conf. Comput. Intell. Virtual Environments Meas. Syst. Appl. (CIVEMSA)*, Jun. 2020, pp. 1–6.

[43] L. Gézero and C. Antunes, "Automated three-dimensional linear elements extraction from mobile LiDAR point clouds in railway environments," *Infrastructures*, vol. 4, no. 3, p. 46, Jul. 2019.

[44] J. Grandio, B. Riveiro, M. Soilán, and P. Arias, "Point cloud semantic segmentation of complex railway environments using deep learning," *Autom. Construct.*, vol. 141, Sep. 2022, Art. no. 104425.

[45] D. Lamas, M. Soilán, J. Grandío, and B. Riveiro, "Automatic point cloud semantic segmentation of complex railway environments," *Remote Sens.*, vol. 13, no. 12, p. 2332, Jun. 2021.

[46] M. Soilán, A. Nóvoa, A. Sánchez-Rodríguez, A. Justo, and B. Riveiro, "Fully automated methodology for the delineation of railway lanes and the generation of IFC alignment models using 3D point cloud data," *Autom. Construct.*, vol. 126, Jun. 2021, Art. no. 103684.

[47] A. Gutiérrez-Fernández, C. Fernández-Llamas, V. Matellán-Olivera, and A. Suárez-González, "Automatic extraction of power cables location in railways using surface LiDAR systems," *Sensors*, vol. 20, no. 21, p. 6222, Oct. 2020.

[48] J. Jung, L. Chen, G. Sohn, C. Luo, and J.-U. Won, "Multi-range conditional random field for classifying railway electrification system objects using mobile laser scanning data," *Remote Sens.*, vol. 8, no. 12, p. 1008, Dec. 2016.

[49] Y. Jwa and G. Sonh, "Kalman filter based railway tracking from mobile LiDAR data," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. II-3/W5, pp. 159–164, Aug. 2015.

[50] A. Karmacharya, F. Boochs, and B. Tietz, "Knowledge guided object detection and identification in 3D point clouds," *Proc. SPIE*, vol. 9528, Jun. 2015, Art. no. 952804.

[51] A. Karunathilake, R. Honma, and Y. Niina, "Self-organized model fitting method for railway structures monitoring using LiDAR point cloud," *Remote Sens.*, vol. 12, no. 22, p. 3702, Nov. 2020.

[52] B. Kwoczynska, W. Sagan, and K. Dziura, "Elaboration and modeling of the railway infrastructure using data from airborne and mobile laser scanning," in *Proc. Baltic Geodetic Congr. (BGC Geomatics)*, Jun. 2016, pp. 106–115.

[53] F. Li, X. Ren, W. Luo, and X. Chen, "Methodology for existing railway reconstruction with constrained optimization based on point cloud data," *Appl. Sci.*, vol. 8, no. 10, p. 1782, Oct. 2018.

[54] S. Lin, C. Xu, L. Chen, S. Li, and X. Tu, "LiDAR point cloud recognition of overhead catenary system with deep learning," *Sensors*, vol. 20, no. 8, p. 2212, Apr. 2020.

[55] S. Liu, X. Tu, C. Xu, L. Chen, S. Lin, and R. Li, "An optimized deep neural network for overhead contact system recognition from LiDAR point clouds," *Remote Sens.*, vol. 13, no. 20, p. 4110, Oct. 2021.

[56] X. Tu, C. Xu, S. Liu, S. Lin, L. Chen, G. Xie, and R. Li, "LiDAR point cloud recognition and visualization with deep learning for overhead contact inspection," *Sensors*, vol. 20, no. 21, p. 6387, Nov. 2020.

[57] A. Mahtani, N. Chouchani, M. Herbreteau, and D. Rafin, "Enhancing autonomous train safety through a priori-map based perception," in *Reliability, Safety, and Security of Railway Systems. Modelling, Analysis, Verification, and Certification* (Lecture Notes in Computer Science), vol. 13294. Cham, Switzerland: Springer, 2022, pp. 115–129.

[58] A. Manier, J. Moras, J.-C. Michelin, and H. Piet-Lahanier, "Railway LiDAR semantic segmentation with axially symmetrical convolutional learning," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 2, pp. 135–142, May 2022.

[59] A. Marwati and C. K. Wang, "Automatic retrieval of railway masts tilt angle from mobile laser scanning data," in *Proc. 42nd Asian Conf. Remote Sens. (ACRS)*. Can Tho, Vietnam: Asian Association on Remote Sensing, 2021.

[60] S. O. Elberink, K. Khoshelham, M. Arastounia, and D. D. Benito, "Rail track detection and modelling in mobile laser scanner data," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 2, no. 5, pp. 223–228, 2013.

[61] E. Pastucha, "Catenary system detection, localization and classification using mobile scanning data," *Remote Sens.*, vol. 8, no. 10, p. 801, Sep. 2016.

[62] S. J. Roberts, L. K. Bonenberg, H. Jing, A. Sowter, X. Meng, T. Moore, C. Hill, and P. Bhatia, "Monitoring rail infrastructure using multisensor navigation on a moving platform and autonomous robots," in *Proc. 30th Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS+)*, vol. 3, Nov. 2017, pp. 1928–1935.

[63] S. Sahebdivani, H. Arefi, and M. Maboudi, "Rail track detection and projection-based 3D modeling from UAV point cloud," *Sensors*, vol. 20, no. 18, p. 5220, Sep. 2020.

[64] M. Soilán, A. Justo, A. Sánchez-Rodríguez, D. Lamas, and B. Riveiro, "3D point cloud data processing and infrastructure information models: Methods and findings from safeway project," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 43, pp. 239–246, Jun. 2021.

[65] G. Uggla and M. Horemuz, "Towards synthesized training data for semantic segmentation of mobile laser scanning point clouds: Generating level crossings from real and synthetic point cloud samples," *Autom. Construct.*, vol. 130, Oct. 2021, Art. no. 103839.

[66] Z. Wang, G. Yu, P. Chen, B. Zhou, and S. Yang, "FarNet: An attention-aggregation network for long-range rail track point cloud segmentation," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 13118–13126, Aug. 2022.

[67] L. Xu, S. Zheng, J. Na, Y. Yang, C. Mu, and D. Shi, "A vehicle-borne mobile mapping system based framework for semantic segmentation and modeling on overhead catenary system using deep learning," *Remote Sens.*, vol. 13, no. 23, p. 4939, Dec. 2021.

[68] B. Yang and L. Fang, "Automated extraction of 3-D railway tracks from mobile laser scanning point clouds," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 12, pp. 4750–4761, Dec. 2014.

[69] X. Yu, W. He, X. Qian, Y. Yang, T. Zhang, and L. Ou, "Real-time rail recognition based on 3D point clouds," *Meas. Sci. Technol.*, vol. 33, no. 10, Oct. 2022, Art. no. 105207.

[70] L. Zhu and J. Hyyppa, "The use of airborne and mobile laser scanning for modeling railway environments in 3D," *Remote Sens.*, vol. 6, no. 4, pp. 3075–3100, Apr. 2014.

[71] R. Zou, X. Fan, C. Qian, W. Ye, P. Zhao, J. Tang, and H. Liu, "An efficient and accurate method for different configurations railway extraction based on mobile laser scanning," *Remote Sens.*, vol. 11, no. 24, p. 2929, Dec. 2019.

[72] S. A. Bello, S. Yu, C. Wang, J. M. Adam, and J. Li, "Deep learning on 3D point clouds," *Remote Sens.*, vol. 12, no. 11, p. 1729, 2020.

[73] A. Nurunnabi, G. West, and D. Belton, "Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point cloud data," *Pattern Recognit.*, vol. 48, no. 4, pp. 1404–1419, Apr. 2015.

[74] Y. Guo, F. Sohel, M. Bennamoun, J. Wan, and M. Lu, "A novel local surface feature for 3D object recognition under clutter and occlusion," *Inf. Sci.*, vol. 293, pp. 196–213, Feb. 2015.

[75] J. Wolf, R. Richter, and J. Döllner, "Asset detection in railroad environments using deep learning-based scanline analysis," in *Proc. 16th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, vol. 4, Setúbal Municipality, Portugal, Jan. 2021, pp. 465–470.

[76] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[77] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.

[78] E. Grilli, F. Menna, and F. Remondino, "A review of point clouds segmentation and classification algorithms," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 42, pp. 339–344, Feb. 2017.

[79] N. Luo, Y. Jiang, and Q. Wang, "Supervoxel-based region growing segmentation for point cloud data," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 35, no. 3, Mar. 2021, Art. no. 2154007.

[80] J. Li, M. Wu, and H. Li, "3D reconstruction and volume estimation of jujube using consumer-grade RGB-depth sensor," *IEEE Access*, vol. 11, pp. 61502–61512, 2023.

[81] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph CNN for learning on point clouds," *ACM Trans. Graph.*, vol. 38, no. 5, pp. 1–12, Oct. 2019.

[82] L. Landrieu and M. Simonovsky, "Large-scale point cloud semantic segmentation with superpoint graphs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* Los Alamitos, CA, USA: IEEE Computer Society, Jun. 2018, pp. 4558–4567.

[83] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, and M. Bennamoun, "Deep learning for 3D point clouds: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4338–4364, Dec. 2021.

[84] G. Riegler, A. O. Ulusoy, and A. Geiger, "OctNet: Learning deep 3D representations at high resolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6620–6629.

[85] H. Lei, N. Akhtar, and A. Mian, "Spherical kernel for efficient graph convolution on 3D point clouds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3664–3680, Oct. 2021.

[86] P.-S. Wang, Y. Liu, Y.-X. Guo, C.-Y. Sun, and X. Tong, "O-CNN: Octree-based convolutional neural networks for 3D shape analysis," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–11, Aug. 2017.

[87] Y. Xu, X. Tong, and U. Stilla, "Voxel-based representation of 3D point clouds: Methods, applications, and its potential use in the construction industry," *Autom. Construct.*, vol. 126, Jun. 2021, Art. no. 103675.

[88] R. Klokov and V. Lempitsky, "Escape from cells: Deep *kd*-networks for the recognition of 3D point cloud models," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 863–872.

[89] H. Thomas, C. R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, and L. Guibas, "KPConv: Flexible and deformable convolution for point clouds," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6410–6419.

[90] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: Enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020.

[91] M. Soilán, A. Sánchez-Rodríguez, P. Del Río-Barral, C. Perez-Collazo, P. Arias, and B. Riveiro, "Review of laser scanning technologies and their applications for road and railway infrastructure monitoring," *Infrastructures*, vol. 4, no. 4, p. 58, Sep. 2019.

[92] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Providence, RI, USA, Jun. 2012, pp. 3354–3361.

[93] B. Ton, "Labelled high resolution point cloud dataset of 15 catenary arches in The Netherlands," 4TU Res. data, The NetherlandsTech. Rep., 2022.

[94] B. Dekker, "Semi-supervised point cloud segmentation on railway data," M.S. thesis, Dept. Comput. Sci., Univ. Twente, Enschede, The Netherlands, 2023.

[95] J. Sauder and B. Sievers, "Self-supervised deep learning on point clouds by reconstructing space," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2019, pp. 1–11.

[96] T. Baltrusaitis, C. Ahuja, and L.-P. Morency, "Multimodal machine learning: A survey and taxonomy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 2, pp. 423–443, Feb. 2019.

[97] Z. Zhang, W. Sun, X. Min, Q. Wang, J. He, Q. Zhou, and G. Zhai, "MM-PCQA: Multi-modal learning for no-reference point cloud quality assessment," in *Proc. 32nd Int. Joint Conf. Artif. Intell.*, Aug. 2023, pp. 1759–1767.

[98] W. Samek, "Explainable deep learning: Concepts, methods, and new developments," in *Explainable Deep Learning AI*, J. Benois-Pineau, R. Bourqui, D. Petkovic, and G. Quénot, Eds. New York, NY, USA: Academic Press, 2023, pp. 7–33.

[99] D. Garrido, R. Rodrigues, A. A. Sousa, J. Jacob, and D. C. Silva, "Point cloud interaction and manipulation in virtual reality," in *Proc. 5th Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Jul. 2021, pp. 15–20.

[100] J. A. Rojas, M. Aguado, P. Vasilopoulou, I. Velitchkov, D. Van Assche, P. Colpaert, and R. Verborgh, "Leveraging semantic technologies for digital interoperability in the European railway domain," in *The Semantic Web—ISWC*, A. Hotho, E. Blomqvist, S. Dietze, A. Fokoue, Y. Ding, P. Barnaghi, A. Haller, M. Dragoni, and H. Alani, Eds. Cham, Switzerland: Springer, 2021, pp. 648–664.

[101] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data—The story so far," *Int. J. Semantic Web Inf. Syst.*, vol. 5, no. 3, pp. 1–22, Jul. 2009.

[102] J. Tutcher, "Ontology-driven data integration for railway asset monitoring applications," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Oct. 2014, pp. 85–95.

BRAM DEKKER received the B.Sc. degree in informatics from the University of Amsterdam and the M.Sc. degree in computer science from the University of Twente. His M.Sc. thesis focused on weakly supervised learning for point cloud data, employing techniques, such as active learning and few-shot learning. The goal of his research was the development of segmentation algorithms to facilitate the creation of digital twins for railway infrastructure.

BRAM TON received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Twente, The Netherlands, in 2012. He is currently pursuing the Ph.D. degree with the Saxion University of Applied Sciences, The Netherlands. Between 2013 and 2017, he worked on several projects ranging from finger vein biometrics, and pressure sensors for the automotive sector, to discrete particle modeling. During this period, he was also a Researcher and a Software Developer at a start-up company. This start-up focused on processing multi-view football match video data to create a digital track record of the match. He is also a Medior Researcher with the Saxion University of Applied Sciences. Within this role, he is working on object detection within point cloud data from the railway scene and incorporating this work into his thesis. His research interests include deep learning, computer vision, point clouds, and mixed reality.

JOANNEKE MEIJER received the B.S. and M.S. degrees in econometrics and operations research from Rijksuniversiteit Groningen, The Netherlands, in 2016. Since 2017, she has been a Data Scientist/Consultant/Manager on several projects ranging from forecasting to optimization and dashboarding to natural language processing (NLP). She has worked within different sectors, such as healthcare, finance, retail, sustainability, and the Dutch public sector. During this period, she was also one of the speakers at ODSC Europe talking about sustainable retail through open source, scraping, and NLP. In 2022, she was a Researcher with the Saxion University of Applied Sciences, The Netherlands. Within this role, she worked on both applied data science projects and augmented interaction projects.

NACIR BOUALI received the M.Sc. degree in software engineering from Al Akhawayn University, Iran, in 2016. He is currently pursuing the Ph.D. degree in virtual reality generation from natural language in collaboration with the University of Eastern Finland. He is a Lecturer with the University of Twente, The Netherlands, where he teaches courses on artificial intelligence, machine learning, and databases. His research interests include AI in education, natural language processing, and the use of immersive virtual environments in language learning.

JEROEN LINSSEN received the Graduate degree in cognitive artificial intelligence from Utrecht University and the Ph.D. degree in human–media interaction from the University of Twente. He is currently a Lector in symbiotic artificial intelligence with the Research Group Ambient Intelligence, Saxion University of Applied Sciences. Data is central to the research performed at the group he co-leads, with research lines on connected embedded systems (the Internet of Things and data acquisition), applied data science (AI and machine learning), and augmented interaction (augmented and virtual reality). His current research interests include explainable AI and MLOps to provide applicable insights into man-machine collaboration, serious games, social robotics, and intelligent agents.

FAIZAN AHMED received the Ph.D. degree in applied mathematics from the University of Twente.

He is currently a Lecturer with the Formal Methods and Tools Group, University of Twente. As a Lecturer, he teaches courses on databases, machine learning, and explainable AI. He also supervises B.Sc., M.Sc., and Ph.D. students, specializing in applied machine learning, particularly focusing on machine learning for point cloud data, time series data, and explainable AI, all with a strong emphasis on practical applications in industry and healthcare. Additionally, he is a part of the Ambient Intelligence Group, Saxion University of Applied Sciences, where he primarily works on applying advanced machine learning techniques to solve industrial problems. He is also actively involved in a project aimed at creating a digital twin for the Dutch railways using point cloud data. His research interests include applied machine learning and explainable AI.