Integration of Building Information Modeling and machine learning for railway defect localization

Jessada Sresakoolchai and Sakdirat Kaewunruen
Department of Civil Engineering, University of Birmingham, Birmingham B15 2TT, United Kingdom

Corresponding author: Sakdirat Kaewunruen (e-mail: kaewunruen@bham.ac.uk).

The authors also wish to thank the European Commission for the financial sponsorship of the H2020-RISE Project no.691135 “RISEN: Rail Infrastructure Systems Engineering Network”, which enables a global research network that addresses the grand challenge of railway infrastructure resilience and advanced sensing in extreme environments (www.risen2rail.eu).

ABSTRACT Building Information Modeling (BIM) has been used in various industries for a long time. The railway system is another industry where BIM plays an important role. Since BIM can contain project information in different stages, a pool of information is involved and included in BIM. To use this information efficiently, machine learning, as a branch of artificial intelligence, is one of the tools widely applied nowadays. However, integrating BIM and machine learning in the railway system is new. This study aims to integrate BIM and machine learning to localize defects in the railway infrastructure. In this study, wheelburns are used as case studies. Machine learning techniques used to localize defects are Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). From the study, the developed BIM model can be fully integrated with machine learning to localize defects in the railway infrastructure using the developed workflow. It is found that the CNN model provides the best outcome when Mean Absolute Error (MAE) is used as the main indicator. The MAE of the CNN model is 0.03 m and the Max Error (ME) is 0.3 m. The results of the study show that the integration of BIM and machine learning can be achieved and provide advantages to the railway industry. The developed machine learning models provide satisfactory performance and will be beneficial for the railway industry for better asset management and cost-effective maintenance.

INDEX TERMS Building Information Modeling, Machine Learning, Artificial Intelligence, Railway Maintenance, Wheelburns, Defect Localisation

I. INTRODUCTION

Building Information Modeling (BIM) is the development of models to simulate a project in different stages from the planning stage to the decommissioning stage. The model contains a pool of information stored in digital form. The advantages of BIM are a more effective process in terms of time and cost, better design, better control, better quality, enhancing automation and collaboration, and better life cycle management [1]. For its advantages, BIM has been used for a long time especially in the architecture, engineering, and construction (AEC) industry. However, BIM has not been used as extensively in the railway system as in the AEC industry but its role there is increasing. Many governments around the world require BIM to be used in public projects. For example, the UK government required BIM Level 2 to be implemented in infrastructure projects by 2016 [2]. In this case, BIM Level 2 represents a 3D model containing information and different parties combined their BIM models for collaboration. Although it can deliver many benefits from its application, BIM has not been used widely in the railway industry as much as in other industries. A function of BIM is to manage information. Without BIM, the railway project and information are not fully integrated. The integration of information and the project can improve the efficiency of project management. In the information management aspect, every stage of the railway project can be integrated with BIM. However, the stage that the railway project most involves with information and where information is the most dynamic factor is the maintenance stage. This information will not be valuable if it is not used. Machine learning can make a benefit from this information; therefore, integration
of BIM and machine learning can improve the overall efficiency of railway project management.

For the railway system, the operation and maintenance stages are the longest stages in the life cycle. During the operation and maintenance stages, a great deal of information emerges. Using BIM to contain information will benefit from the information such as in defect detection or defect localization. Due to the increasing demand for railway transportation, rolling stock has to operate faster and loads increase so the deterioration accelerates.

In this study, wheelburns, which are common defects in rails, are used as case studies. Sawley and Reiff [3] collected railway defect information around the world and found that wheelburns account for about 13–17% of all defects. Wheelburns are caused by the slipping of wheels and rails. Examples of causes of wheelburns are excessive grades, poor driving procedures, insufficient locomotive power, and contamination [4]. If wheelburns are not detected in time, they will cause more serious defects and accelerate other defects, resulting in serious incidents.

To detect railway defects, machine learning and other techniques are applied such as laser [5], acoustic emission [6], or image processing [7]. It has been proven that machine learning is capable of detecting defects and the performance is satisfied [8].

This study presents a defect localization approach using machine learning techniques. Defect inspection can be done using different techniques such as visual inspection, ultrasonic technology, laser or other optical techniques, and computer image processing. However, each technique has some limitations. For example, visual inspection requires time and manpower to conduct, ultrasonic techniques require special equipment to be installed, or computer image processing has a limitation of light. Applying an inertia technique tends to be more interested because it requires little additional installation for collecting data. Installing acceleration sensors is enough to measure. Moreover, data collection is relatively fast compared to other techniques. This study applied Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) to develop models for defect localization using axle box accelerations. Data used in this study is simulated using D-Track which is verified software for simulating the dynamic behavior of rolling stock and rails.

This study aims to integrate BIM and machine learning to localize defects in the railway infrastructure and wheelburns are selected as case studies. If BIM and machine learning are fully integrated, the management will be easier, more suitable, and more reasonable because decision-making is executed based on information. The time requiring humans to inspect, diagnose, and respond will thereby be reduced. Therefore, defects can be managed in time and properly. The study is expected to demonstrate the potential of the developed workflow for integrating BIM and machine learning which will be beneficial for the railway industries in different aspects.

II. LITERATURE REVIEW

A. BUILDING INFORMATION MODELING (BIM)

Building Information Modeling is a virtual model that is digitally created. It contains project information from the beginning to the end of the project [9]. Related parties can apply BIM for the benefit of the project at different stages. Besides supporting project visualization, one of the main benefits of BIM is supporting integration and collaboration. It can be classified as levels; according to the UK government’s 2011 Construction Strategy, there are four levels of BIM [2]. Level 0 indicates that the model is done using 2D computer-aided design (CAD). The collaboration is limited and information exchange is done using paper. Level 1 indicates that 2D and 3D models are used. Information exchange is done electronically. Level 2, which the UK government aimed to achieve in 2016, indicates that the model is developed as a 3D model containing information. Relevant parties can combine their models for collaboration and sharing. Finally, level 3 indicates that a single model is used in the project for information exchange. Today, BIM level 2 is a standard for public and private projects around the world and BIM level 3 is an ultimate goal that everyone aspires to.

Building Information Modeling can also be classified using its dimensions. A BIM model is created on a 3D basis so information on shapes and components is included in 3D models. However, BIM models are further developed to 4D, 5D, or 6D dimensions to fulfill more objectives. Each dimension has its own purpose. For example, 4D is used for planning by representing time or schedule, 5D is used for determining cost or quantity, and 6D can be used for the sustainability purpose by considering energy consumption or carbon emissions [10]. It can be noted that the developed BIM model in this study is 6D BIM because project scheduling and budget, which are 4D and 5D, can be considered by the software used to generate the BIM model. At the same time, integrating BIM and machine learning for defect localization is 6D for maintenance purposes.

In the railway system, BIM plays a more important role as mentioned, and BIM was applied with Geographic Information systems (GIS). The integration was mainly used for project planning and decision-making [11]. However, the guideline for collaboration was not clear and needed to be developed further. In 2018, Bensalah et al. [12] implemented BIM in a railway project in Morocco. The results from their study confirmed the advantages and disadvantages of implementing BIM according to the literature. Shin et al. [13] also analyzed the benefit of applying BIM in railway projects. They investigated 7 railway projects in South Korea and found that 12 errors could be prevented by using BIM. From the analysis, applying BIM could save more than $50,000 when comparing the costs of BIM provision and
error fixing. As a specific use of BIM in the railway system, BIM was used for the sustainability evaluation of railway stations [14]. The 6D version of the BIM model in that study was used to calculate carbon emissions when stations were renovated or reconstructed and select the optimal option. It can also be used in the railway system for other purposes such as turnouts’ lifecycle management [15] or vulnerability audits for subway stations [16]. From the literature, we conclude that using BIM for railway maintenance purposes is still limited in scope and application.

B. DEFECT DETECTION AND LOCALIZATION

The maintenance cost is responsible for the highest proportion of the railway life cycle cost [17], hence, better maintenance can reduce that significantly. For the maintenance cost, 20–40% of the budget is spent on the inspection function [18]. Therefore, a good approach for detecting and localizing defects can result in lowering the cost of the railway project. Railway defects have to be managed properly for safety, passenger comfort, and cost-effectiveness. A capability to detect defects in time is important. For example, small wheelburns can be treated by rail grinding, the cost of which is significantly lower than track replacement [19]. Railway defects can be detected using many techniques such as site inspection, acoustic emission [6], image processing [7], laser [5], electromagnetic tomography [20], machine vision [21]. One technique which is popular nowadays is that of machine learning. In 2017, Krummenacher et al. [22] proposed a method to detect defects by installing sensors on tracks and using wheel vertical force as the main features. They used the Support Vector Machine (SVM) and CNN to detect flat spots, shelling, and non-roundness. The accuracy of their models was 80–90%, and CNN was also used to detect defects in insulators as image processing with the accuracy of more than 90% [23] as well as surface defects [24]. From the literature, detecting defects using machine learning is usually conducted by image processing which requires additional installation and has some limitations such as that the system cannot be used when the light is not sufficient or in an area without cameras. The authors proposed an approach using accelerations that could be collected from axle box acceleration sensors [8] and the predictive model was developed by CNN using time-series data. The advantages of this approach were that additional installations were not required, there was no limitation about the light needed, data could be collected continuously because sensors are on board, and the prediction was fast. The results showed that the accuracy was more than 95% on detection and severity classification.

As to defect detection, defect localization can be conducted using traditional methods such as site inspection or sensors. However, machine learning is one of the interesting techniques that can be used as well. Railway defect localization is limited in terms of research while defect localization has been of interest in other areas. In 2004, Poyhonen et al. [25] found that using SVM provided the best outcome when used to localize insulation defects in electrodes. Ferguson et al. [26] applied CNN to detect and localize casting defects in metal casting using images as input. The precision of CNN was about 90%. From the literature, studies about defect localization in the railway system are scant. From the literature review, there is a research gap in BIM and machine learning integration and railway defect localization. This study, therefore, aims to fill this gap by presenting the workflow to integrate BIM and machine learning which is used to localize railway defects.

III. METHODOLOGY

A. BIM MODEL DEVELOPMENT

The first step of this study is to develop a BIM model. In this study, the BIM model represents a route of a double-track railway project. The BIM model is developed using AutoCAD Civil 3D and aims to be a 6D BIM model. The 4D, 5D, and 6D models in this study refer to the time, cost, and maintenance aspects, respectively. The process of development can be described as follows.

The 3D BIM model is developed based on vertical and horizontal alignments. These include all necessary components of the railway line such as point of intersection, superelevation, a radius of curvature, or gradient. The design has to conform to the defined standard. The software provides functions to detect these compatibilities. For example, if the radius of curvature is too small for the design speed, the software will show a notification. In this step, sample lines are defined to represent sections of the line. In this study, sample lines are defined every 20 m so the length of each section in the BIM model is 20 m. After the vertical and horizontal alignments are designed, the next step, which is a crucial step, is assembly definition. In this study, a cross-section of the railway line is defined. Rail components such as rails, sleepers, embankment, foundation, retaining wall, or tunnels are defined in this step. The next step is corridor creation. In this step, vertical alignment, horizontal alignment, and assembly are combined to create the corridor as a 3D solid. The 3D solid is exported as an Industry Foundation Classes (IFC) file which is exchangeable and compatible with other BIM software. In this step, the 3D BIM model is created. In this study, a traditional double track and railway tunnel are demonstrated in Figure 1.
To develop the BIM model to 4D, project scheduling software is applied. In this study, Navisworks is used to link railway components to the project schedule. Any railway components can be scheduled depending on the desired level of detail (LOD). The model is appended to Navisworks via the exported IFC file. The developed schedule can be managed using popular project management software such as Microsoft Project or Primavera P6. The progress of the project can be tracked and the simulation of the project can be made. The 5D BIM model is developed based on the 4D model. The assembly which is defined in the 3D creation process is used to calculate the quantity and cost of the project. The quantity is calculated based on rail components and terrain. For example, the length of rails is calculated based on the alignment and soil work is calculated based on the alignment and terrain. Civil 3D has a function to calculate this quantity. Calculation methods are various and can be selected based on preference. This can be done using the schedule function provided in Civil 3D. It is to be noted that the schedule in Civil 3D refers to the tabular function, and not the time management.

The 6D BIM model is developed for maintenance purposes. The operational characteristics are contained in the model in this step. Examples of information contained are design speed, the weight of rolling stock, and acceleration. This information is stored in the model via Property Set Definitions. A new Property Definition can be added manually or automatically; however, the automatic approach is more effective and needs support from other software. One-value Property Definitions can be contained in the model; however, multiple-value or time-series Property Definitions are not suitable to be contained in the BIM model because they increase the size of the model unnecessarily and slow down the process of the work. Therefore, this study proposes that time-series Property Definitions should be contained in the model as hyperlinks instead. From all steps, the 6D BIM model is created and ready to integrate with machine learning.

**B. DATA PREPARATION**

In this study, data is numerically generated by simulations. The software used for simulations is D-Track which is a dynamic track behavior simulation, developed by Cai [27] in 1996. However, there was an issue with the accuracy compared to the field data. Steffens [28] then integrated the user interface and Dynamic Analysis of Rail Track Structure (DARTS) in D-Track developed by Cai but the difference between field data and simulation still existed. Leong [29] improved D-Track based on Australian Standard code AS1085.14. He finally found the causes that created the issues in D-Track, namely too-low wheel-rail force calculation, unnecessary assumptions in D-track, improper sleeper pad reactions, and an inaccurate sleeper bending moment calculation. After improvements, the accuracy of the software is more than 90%, therefore, it can be concluded that results from D-Track simulations are reliable and accurate. Also, it ensures that the simulations can be referred to as the real site inspection data. Although data from simulations is free from noise, site data can also be processed to filter the noise before being used to train machine learning models. From this approach, it can be inferred that outcome from simulations is validated to represent site inspection data.

To prepare data for defect localization, wheelburns are used and track profiles are created to refer to the differences in wheelburns’ size and location. Other parameters that varied to create data variation are the weights and speeds of the rolling stock. The total number of simulations is 1,620. The simulation summary is shown here:

| Parameters            | Variation                        |
|-----------------------|----------------------------------|
| Type of defect        | Wheelburns                       |
| Sizes of defect       | 25, 50, and 75 mm referred to Australian Rail Track Corporation [4] |
| Locations of defect   | 0-20 m                           |
| Weight of rolling stock| 40, 60, and 80 tons              |
| Speed of rolling stock| 20-200 km/h                      |

D-Track can generate varied results such as wheel–rail contact forces, axle accelerations, sleeper shears, or sleeper bending moments. In this study, axle accelerations are used for the aforementioned reasons. An example of axle accelerations from D-Track is shown in Figure 2. From this figure, the parameters are the following: the size of wheelburns is 25 mm, the location of wheelburns is 11.5m, the weight of rolling stock is 80 tons, and the speed of rolling stock is 200 km/h. Other parameters are defaulted by the software. It can be seen that the data may look simple however when the speed of rolling stocks change, locations of defects are difficult to localize because the input is based on time not location as shown in Figure 2. The number of data points is 6,695 in a simulation. The accelerations are stored in the BIM model as hyperparameters because they are time-series data that are not suitable to be stored in the BIM model. Accelarations are used as features for defect localization.
C. BIM AND MACHINE LEARNING INTEGRATION AND INFORMATION MANAGEMENT

The workflow used to integrate BIM and machine learning is developed as shown in Figure 3. From the figure, the workflow starts with inspection data. This data can be collected using installed sensors or site inspection. Then, data is transformed for the appropriate storage and application. Collected data can be stored in servers, the cloud, or a local repository. In this study, the local repository is used to demonstrate. Until this stage, data is ready to be synced with the BIM model. Before including data into the BIM model, the BIM model has to be designed first with which Property Definitions are required to store in the BIM model. In this study, the weight of rolling stock, speed of rolling stock, and axle accelerations are stored in the BIM model. Other Property Definitions can also be stored in the BIM model depending on the requirement of use.

When the BIM model is designed and Property Definitions are set, the BIM model is ready to exchange information with other platforms. To store information in the BIM model, Dynamo is used. Dynamo is a visual programming environment developed by AutoCAD, which is used to manage and automate the BIM model. The Property Set Definitions are exported using Dynamo. Then, a spreadsheet file that contains all Property Definitions is exported. Visual Basic for Applications (VBA) is used to manage the spreadsheet file. Inspection data and other information are included in the exported file from the BIM model, thus, the exported file contained all the necessary information. This file is updated Property Set Definitions which can be updated to the BIM model using Dynamo.

Returning to the exported spreadsheet file containing Property Set Definitions, data is extracted to develop machine learning models. Simplified data and raw data can be extracted from the exported spreadsheet file. In this case, raw data is extracted to localize defects. Data extraction is conducted using VBA to prepare the input for training machine learning models. In the application aspect, data used to make the prediction is prepared in this step. Then, machine learning models are trained or the prediction is made. The outcome is machine learning models or predictions which are

Figure 2. An example of accelerations simulated by D-Track

Figure 3. Workflow for integrating BIM and machine learning

To exchange information between the BIM model and the spreadsheet file, Dynamo is used with Python script which is supported in Dynamo. As mentioned, one-value data can be stored in the BIM model. However, accelerations which are time-series data are not suitable. This study proposes to store this class of data in the form of hyperlinks. To include the hyperlinks in the BIM model, Dynamo is applied which is shown in Figure 4. Now, the BIM model contains maintenance information and characteristics of rolling stock which are ready to be used to develop the machine learning models as shown in Figure 5.

Figure 4. Dynamo used for including hyperlinks in the BIM model

Figure 5. An Example of Property Sets when data is included
stored in the BIM model using Dynamo and VBA as mentioned.

D. MACHINE LEARNING MODEL DEVELOPMENT

All types of neural networks—DNN, CNN, and RNN—are used to develop predictive models for defect localization. Reasons for using these techniques are they are suitable for problems with non-linear characteristics, they are suitable for time-series data, and they are proof that they are effective when dealing with these problems. A problem for the machine learning models in this study is the regression problem because predictive values are continuous. An initial shape of input is 1*6,695 because a sample contains 6,695 data points which are continuous values. For the DNN model, the number of input nodes is fixed to 6,695 or equal to the number of data points. The number of output nodes is 1 because the problem is a regression for every model in this study.

The main characteristic of the CNN model is the feature extraction element. This feature makes CNN stand out because it does not require the feature engineering knowledge to select features for developing models; the machine will do it. Therefore, if the model is developed properly, it chooses the best set of features for the prediction. The feature extraction part of the CNN model consists of convolution layers while pooling layers and dropouts are optional. Pooling layers are used to minimize the size of the feature map which makes the training faster, while dropouts are used to prevent overfitting. Another feature of the CNN model is the classification or regression element and in this study, it is the regression element. The characteristic of the regression element is the same as the DNN model—it is constructed from fully-connected or dense layers.

Unlike other techniques, RNN has internal memories. This allows the model to explore the relationship between features rather than considering each feature independently like other techniques. Therefore, RNN is popular in handwriting or speech recognition where the sequence of words is important. In this study, Long Short-Term Memory (LSTM) which is the improved RNN, is used. It does not have a vanishing gradient problem as with the traditional RNN; as the CNN model, the LSTM part connects to the regression element.

Mean Absolute Error (MAE) is used to indicate the performance of the models. MAE is used rather than other indicators such as Root Mean Square Error (RMSE) or Mean Absolute Percentage Error (MAPE) because it is easy to understand, straightforward, and corresponds to the value to be measured. Of the data, 70% is used as training data and the remaining 30% is testing data.

To make sure the models provide the best performance, model tuning is required; however, not all parameters are tuned during the training. Those parameters are called hyperparameters, and to perform hyperparameter tuning, different techniques can be used. In this study, Grid Search is used. The list of hyperparameters of each model is shown in Error! Reference source not found..
The performance of each model is assessed by comparing actual data and the prediction. Because numerical labels in this study are locations, not amounts, the accuracy is not suitable to be used to measure the performance of the models. While RMSE is popularly used to assess the performance in regression problems, errors are squared before being averaged. Therefore, the interpretation might not be accurate in this study so MAE is used instead. From Figure 6, it can be seen that DNN is the best model for defect localization because of the MAE where the main indicator is 0.03 m or 3 cm. Compared to the section length of 20 m, the MAE is relatively small or 99.85%. At the same time, the ME is 0.3 m which is acceptable. Compared to other techniques, the accuracy of laser is 0.003 m [30] and the accuracy of ultrasonic is 97%-99% [31]. The performance of the proposed model is acceptable. The MAEs of DNN and RNN are close where DNN provides a slightly smaller MAE than RNN. However, the ME of RNN is significantly larger. Therefore, RNN is not suitable for defect localization in this study. Also, the training time of RNN is significantly higher than DNN and CNN. Therefore, it can be concluded that RNN is not suitable in this case. Although the ME of DNN is smaller than RNN, it is still significantly higher than CNN’s ME. Therefore, the CNN model is the best for the localization of wheelburns.

For the discussion, CNN performs the best because the feature is time-series data whose benefit from feature extraction is outstanding, as CNN can recognize the pattern in time-series data and localize defects accurately. Moreover, the machine learns to extract suitable features from the training. Some insignificant features are ignored during the training via feature extraction and pooling layer; therefore, the prediction is done with low error. For DNN, all features or 6,695 data points are used as input and then they are passed to hidden layers. Although significant and insignificant features are categorized via weight and bias, all features are still used for the prediction. This is the main difference between DNN and CNN and therefore, the performance of DNN is not as good as CNN. For RNN, although the MAE is close to CNN, the ME is unacceptably high. Thus, it can be inferred that RNN cannot extract the appropriate features from input to make the prediction because the application is not designed for this kind of task. As mentioned, RNN provides a satisfying outcome when the sequence of features affects the prediction. Different from the sequence of words in sentences, each value of axle accelerations in this study tends to be more independent. Therefore, RNN does not perform as well as other models.

V. DEMONSTRATION
The workflow developed for integrating BIM and machine learning to localize defects is applicable. Based on Figure 3, data is collected using axle box acceleration sensors. Then, it is processed and stored in the selected platform such as a server or cloud system. The processed data is added into the BIM model using Dynamo and Python script which is also used to export data from the BIM model. Exported data is processed to a form that is ready for being trained by machine learning. VBA plays an important role in this step. Data for machine learning is extracted using VBA. Extracted data is used both to train machine learning models and make prediction. Then, the prediction is integrated into the BIM model using Dynamo and Python script to notify officers to take action. For example, machine learning models can detect and localize defects. Then, officers can plan the maintenance. In the beginning, data collection has to be done in parallel with the traditional approach to collect labels for
machine learning models’ training. When the size of data is big enough, the detection or prediction can rely on sensors without labeling. In real practice, when data are collected, they can be used to develop machine learning models. When the models are well-tuned, they can be used to detect and localize defects. For example, axle box accelerations are used to collect vibrations when rolling stocks are operated. Then, vibrations are used as raw data to feed into machine learning models to detect locations of defects. Then, operators know how to respond to defects and where defects are. This approach will increase the efficiency of railway maintenance in terms of time, cost, and quality.

VI. CONCLUSION

This study presents the workflow for integrating BIM and machine learning for railway defect localization. The finding of the study is original and applicable. Wheelburns are selected because they are among the most common defects in the railway system. The numerical data is simulated using D-Track which is the verified railway dynamic behavior simulation. The developed BIM model is designed to be the 6D BIM model. The model is developed using AutoCAD Civil 3D. To be compatible with other BIM software, the BIM model is generated as a 3D Solid which can be exported as an IFC file that is the standard format for BIM models. For information management and exchange, Dynamo and VBA are used to integrate BIM and machine learning. The developed workflow provides seamless integration between BIM and machine learning. The work process is automatic and takes a short time to integrate BIM and machine learning and even time-series data is indirectly stored in the BIM model via hyperlinks. The stored information is beneficial for machine learning and maintenance aspects. The integration extends the benefits of BIM from the 3D model to the information management platform which contains information and used the contained information to manage the railway project.

The neural networks DNN, CNN, and RNN are used to develop machine learning models for wheelburn defect localization. Axle accelerations are used as features for developing predictive models. Hyperparameter tuning is performed to make the models provide the best outcome. The number of samples is 1,620 which are numerically generated using D-Track. Of the samples, 70% are used as training data and 30% are used as testing data. From the machine learning model development, CNN performs the best with MAE of 0.03 m and ME of 0.3 which are lower than DNN and RNN. Moreover, the training time is 1 second/epoch approximately which is slightly longer than DNN. Therefore, it can be concluded that CNN is the best model for defect localization in this study in both performance and cost aspects. Compared to site inspection and other methods, applying machine learning takes significantly less time, is safer for officers, and requires less installation cost. If more data is available in the future, machine learning can be integrated with BIM for benefits in more aspects, and decision-making can be made optimal. The concept is applicable as the explanation in the demonstration part.

For further development, other functions of machine learning can be included in the BIM model such as detection and severity classification. Real data can also improve the practical use of the railway system. Other defects can be included in the consideration. Other forms of information can be tried to develop machine learning models such as images, noise, or signals from any sensors.

ACKNOWLEDGMENT

The authors also wish to thank the European Commission for the financial sponsorship of the H2020-RISE Project no.691135 “RISEN: Rail Infrastructure Systems Engineering Network”, which enables a global research network that addresses the grand challenge of railway infrastructure resilience and advanced sensing in extreme environments (www.risen2rail.eu).

REFERENCES

[1] C. R. C. Construction Innovation, "Adopting BIM for facilities management: Solutions for managing the Sydney Opera House," Cooperative Research Center for Construction Innovation, Brisbane, Australia, 2007.

[2] Infrastructure and Projects Authority. Creating a Digital Built Britain: what you need to know. 2017. Available from: https://www.gov.uk/guidance/creating-a-digital-built-britain-what-you-need-to-know#what-digital-built-britain-aims-to-do.

[3] Sawley, K. and R.P. Reiff, "Rail Failure Assessment for the Office of the Rail Regulator: An Assessment of Railtrack's Methods for Managing Broken and Defective Rails", 2000: Transportation Technology Center.

[4] Australian Rail Track Corporation, "Rail Defects Handbook". 2006.

[5] Xiong, Z., Q. Li, Q. Mao, and Q. Zou, "A 3D laser profiling system for rail surface defect detection," Sensors, 2017. 17(8): p. 1791.

[6] Zhang, X., N. Feng, Y. Wang, and Y. Shen, "Acoustic emission detection of rail defect based on wavelet transform and Shannon entropy," J. Sound Vib., 2015. 399: p. 419-432.

[7] Tastimur, C., H. Yetis, M. Karaköse, and E. Akin, "Rail defect detection and classification with real time image processing technique," IJCSSE, 2016. 5(12): p. 283.

[8] Sresakoolchai, J. and S. Kaewunruen, "Wheelflat detection and severity classification using deep learning techniques," Insight-Non-Destructive Testing and Condition Monitoring, 2021.

[9] Aghbar, S., "Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry," Leader Manag Eng, 2011. 11(3): p. 241-252.

[10] Kaewunruen, S., J. Sresakoolchai, and Z. Zhou, "Sustainability-Based Lifecycle Management for Bridge Infrastructure Using 6D BIM," Sustainability, 2020. 12(6): p. 2436.

[11] Kurvi, S., P. Demian, and T.M. Hassan. Integrating BIM and GIS in railway projects: A critical review, 2017. Association of Researchers in Construction Management (ARCOM).

[12] Bensalah, M., A. Elouadi, and H. Mharzi, "BIM integration to railway projects-case study," Mngnt, 2018. 2: p. 3.

[13] Shin, M.H., H.K. Lee, and H.Y. Kim, "Benefit–Cost Analysis of Building Information Modeling (BIM) in a Railway Site," Sustainability, 2018. 10(11): p. 4303.

[14] Kaewunruen, S. and N. Xu, "Digital twin for sustainability evaluation of railway station buildings," Front. Built Environ., 2018. 4: p. 77.

[15] Kaewunruen, S. and Q. Lian, "Digital twin aided sustainability-based lifecycle management for railway turnout systems," J. Clean. Prod., 2019. 228: p. 1537-1551.
[16] Kaewunruen, S., S. Peng, and O. Phil-Ebosie, "Digital twin aided sustainability and vulnerability audit for subway stations," *Sustainability*, 2020. 12(19): p. 7873.

[17] The Railway Technical Website. A WINDOW ON THE WORLD OF RAILWAY SYSTEMS, TECHNOLOGIES AND OPERATIONS 2019; Available from: http://www.railway-technical.com/trains/train-maintenance/

[18] Luică, P. The high maintenance cost of railway track equipment imposes new solutions. 2011; Available from: https://www.railwaypro.com/wp/the-high-maintenance-cost-of-railway-track-equipment-imposes-new-solutions/.

[19] Fay, M., "Rail Grinding". 2007.

[20] Liu, Z., W. Li, F. Xue, J. Xiafang, B. Bu, and Z. Yi, "Electromagnetic Tomography Rail Defect Inspection," *IEEE Transactions on Magnetics*, 2015. 51(10): p. 1-7.

[21] Cheng, Y., D. HongGui, and F. YuXin. Effects of Faster Region-based Convolutional Neural Network on the Detection Efficiency of Rail Defects under Machine Vision. in 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), 2020.

[22] Krummenacher, G., C.S. Ong, S. Koller, S. Kobayashi, and J.M. Buhmann, "Wheel defect detection with machine learning," *IEEE Trans Intell Transp Syst*, 2017. 19(4): p. 1176-1187.

[23] Kang, G., S. Gao, L. Yu, and D. Zhang, "Deep architecture for high-speed railway insulator surface defect detection: Denoising autoencoder with multitask learning," *IEEE T INSTRUM MEAS*, 2018. 68(8): p. 2679-2690.

[24] Shang, L., Q. Yang, J. Wang, S. Li, and W. Lei. Detection of rail surface defects based on CNN image recognition and classification. 2018. IEEE.

[25] Poyhonen, S., M. Conti, A. Cavallini, G.C. Montanari, and F. Filippetti. Insulation defect localization through partial discharge measurements and numerical classification. 2004. IEEE.

[26] Ferguson, M., R. Ak, Y.-T.T. Lee, and K.H. Law. Automatic localization of casting defects with convolutional neural networks. 2005. IEEE.

[27] Cai, Z., "Modelling of rail track dynamics and wheel/rail interaction." 1994.

[28] Steffens, D.M., "Identification and development of a model of railway track dynamic behaviour." 2005.

[29] Leong, J., "Development of a limit state design methodology for railway track," 2007.

[30] ENGINEER SUPPLY. HOW ACCURATE ARE LASER DISTANCE MEASUREMENTS TOOLS? 2020; Available from: https://www.engineersupply.com/How-Accurate-Are-Laser-Distance-Measurements-Tools.aspx.

[31] Senix Distance and Level Sensors. Ultrasonic Sensor Accuracy. 2020; Available from: https://senix.com/ultrasonic-sensor-accuracy/.