On-Demand Monitoring of Construction Projects through a Game-Like Hybrid Application of BIM and Machine Learning

Farzad Pour Rahimian[a,b,*], Saleh Seyedzadeh[c,d], Stephen Oliver[c,d], Sergio Rodriguez[a], Nashwan Dawood[a]

[a]*School of Computing, Engineering & Digital Technologies, Teesside University, Tees Valley, Middlesbrough, UK.*
[b]*Department of Civil and Environmental Engineering (DICeA), University of Florence, Florence, Italy.*
[c]*Faculty of Engineering, University of Strathclyde, Glasgow G1 1XW, UK.*
[d]*arbnco, Glasgow, UK.*

Abstract

Whilst unavoidable, inspections, progress monitoring, and comparing as-planned with as-built conditions in construction projects do not readily add tangible intrinsic value to the end-users. In large-scale construction projects, the process of monitoring the implementation of every single part of buildings and reflecting them on the BIM models can become highly labour intensive and error-prone, due to the vast amount of data produced in the form of schedules, reports and photo logs. In order to address the mentioned methodological and technical gap, this paper presents a framework and a proof of concept prototype for on-demand automated simulation of construction projects, integrating some cutting edge IT solutions, namely image processing, machine learning, BIM and Virtual Reality. This study utilised the Unity game engine to integrate data from the original BIM models and the as-built images, which were processed via various computer vision techniques. These methods include object recognition and semantic segmentation for identifying different structural elements through supervised training in order to superimpose the real world images on the as-planned model. The proposed framework leads to an automated update of the 3D virtual environment with states of the construction site. This framework empowers project

*Email: f.rahimian@tees.ac.uk

managers and stockholders with an advanced decision-making tool, highlighting the inconsistencies in an effective manner. This paper contributes to body knowledge by providing a technical exemplar for the integration of ML and image processing approaches with immersive and interactive BIM interfaces, the algorithms and program codes of which can help replicability of these approaches by other scholars.

Keywords: Construction Management, Progress Monitoring, Building Information Modelling, Image Processing, Virtual Reality, Machine learning

---

1. Introduction

The complexity of the construction projects and the individualistic approach to every building project result in delays and errors. Majority of construction management and monitoring processes are still conducted traditionally with the use of 2D drawings, reports, schedules and photo logs, making the process complicated and inefficient. The impact of Information and Communication Technology (ICT) integration on progressive improvement in the construction is undeniable [1, 2, 3]. Different technologies and systems have been recently implemented on construction sites to improve project communication, coordination, planning and monitoring, including web-based technologies, cloud computing, Building Information Modelling (BIM), and tracking technologies [4, 5]. These novel applications are usually used in different technological combinations to improve the construction monitoring and allow for comparison of the as-planned and as-built models [6]. BIM capabilities are no longer limited to geometry representation (i.e. 3D virtual objects), and enhance many other aspects of construction projects, such as information management (through semantically rich models), inbuilt intelligence and analysis (via active knowledge-based systems and simulation), as well as collaboration and integration (through digital data exchange) [7].

BIM has been widely used in construction projects for improving communication among various parties during different phases of design and project delivery [8, 9]. However, due to the myriad of issues such as the unpredictable pace of works, continually changing site environments, and the need for constant synchronisations, BIM use has been hindered in the monitoring of construction projects. This is despite the fact that adopting on-demand data acquisition techniques in conjunction with BIM models to compare the state of the sites with the as-planned models has been suggested by many researchers [10, 11, 12, 13], because of the time-consuming and error-prone nature of these activities. Nonetheless, the interoperability between the tools and systems has been identified as the main obstacle, which is presently hard to overcome [14].

For addressing this methodological and technical gap, this study developed a framework and a proof-of-concept prototype, facilitating bilateral coordination of information flow between construction sites and the BIM models. In this framework, computer vision and machine learning (ML) techniques are proposed to help prepare and compose site photographs with the as-planned BIM models. The interoperability and integration of these techniques are facilitated by the aid of Virtual Reality (VR) game engines, such as the Unity engine. The proposed hybrid application of image processing and BIM is expected to enable facilitation of on-demand as-built model update for construction progress tracking. The integration performed in the VR environment (VRE) with the use of a game engine is to enable users to actively participate in the progress evaluation as well as highlighting and reporting the inconsistencies

This paper first outlines an overview of existing technologies for automated construction monitoring, focusing on image processing and visualisation methods. Then, the proposed framework for integration of BIM, ML and VR is discussed, followed by the details of prototype design, applied techniques and algorithms, development strategies, and system architecture of the game like VRE. Finally, the conclusion is presented by wrapping up the work done thus far and describing future research opportunities for improvement of the developed system.

2. Related Studies

2.1. Overview of Real-Time Data Collection Technologies

The data collection technologies for construction project monitoring can be divided into three main categories: a) enhanced information technologies, such as multimedia, emails, voice, and hand-held computing, b) geospatial technologies, like Geographic Information System (GIS), Global Positioning System (GPS), barcoding and Quick Response (QR) coding, Radio Frequency Identification (RFID) and Ultra Wide Band (UWB) tags, and c) image-based technologies, including photogrammetry, videogrammetry and laser scanning [15].

GPS and GIS are commonly used automated asset tracking systems, which can be used for the analysis of construction site equipment operations [16]. Location information of the construction elements is used to identify and track the equipment activity and compile the safety information to improve the decision-making and site management processes [17]. When GIS is integrated with BIM, it can help construction managers identify the best spaces for tower cranes and represent the material progress in supply chain management [18].

Barcoding allows for product identification and is considered as one of the most cost-effective construction tool monitoring methods. However, it is time-consuming, due to the required reading to the tag proximity, and there is a restricted quantity of information contained in each barcode label. Barcode technology is also considered to be unreliable as the tags are prone to damage in the harsh construction environment, or can be lost [15]. QR, on the other hand, provides more information, and it is commonly exploited due to the increasing popularity of mobile phones and tablets utilisation onsite with the available QR code reading applications. The implementation of QR codes has been suggested in conjunction with BIM technology to improve communication on-site and to increase health and safety [19].

RFID incorporates the use of radiofrequency waves instead of light waves, allowing for overcoming the distance issue [15]. This technology has widely been used in conjunction with BIM for leveraging control and monitoring of construction projects [20] as well as productivity optimisation and improving health and safety [21]. UWB signals can be reliable, even beyond walls or any other obstruction that may be faced and are known as one of the most accurate systems for distance positioning, monitoring and tracking, consuming considerably low energy [22].

Laser scanning technology allows for 3D object scanning and data collection for existing environments. The data is generally stored in the form of a 3D point cloud, which can be used for the creation of a digital twin of a building [23, 24]. The primary factors that make the scan-to-BIM method less practical are the required expensive equipment and the time-consuming process for data collection and BIM model creation, which itself is error-prone [25]. Although this method is mainly used for documentation and renovation projects, there is a potentiality for automating construction progress monitoring, which can be achieved through a combination of 3D point cloud acquisition and three-dimensional object recognition to categorise construction elements and 3D models associated with construction schedule [26]. Further to laser scanning, light detection and ranging (LiDAR) technology has been proposed for monitoring construction projects by employing robotics and creating 3D modelling [27].

All these cutting-edge technologies offer several benefits, yet pose specific drawbacks, specifically in supporting monitoring of construction sites for inconsistencies [28]. However, capturing site images is still considered as the easiest and least labour intensive method of gathering information from construction sites [15]. The advent of new photo shooting technologies, such as cheap camera drones and depth image cameras has further facilitated and promoted the construction scene recording actions. The next section is dedicated to reviewing technologies processing construction site photos for monitoring purposes.

2.2. Machine Learning-Based Image Processing for Construction Progress Monitoring

ML techniques found their applications in buiding energy filed early on 1990's [29], however, their use in the field of construction monitoring is very new, and their advantage has not been fully exploited. Automatic image-based modelling techniques for progress monitoring and defects detention has been in the area of interest of many researchers, in both construction and computer science domains [30, 31, 32]. Integration of as-built photographs with 4D modelling using time-lapsed photos was proposed for construction progress monitoring [33]. Later, the same group [34] suggested an automated monitoring system for employing daily construction images and an Industry Foundation Classes (IFC) model-based BIM. In their study, visualisation was achieved through the creation of a 4D as-built model from point cloud images and performing an image classification for progress detection and an as-planned model from BIM. This work was then followed by [35], performing an automated comparison of the actual state and planed state of construction through photometric methods, in order to detect discrepancies and adjustment of the construction schedules.

Kim *et al.* [36] proposed a methodology based on image processing for automated update of 4D models via incorporating (Red-Green-Blue) RGB colour image acquisition, in accordance with specific instructions. The progress identification was also performed by applying 3D-CAD based image filters [37]. Roh *et al.* [11] integrated as-planned BIM models with as-built projects data extracted from the site photographs, which was then overlaid in a 3D walkthrough environment to help estimate the delays in the construction progress.

As manually site photography is time-consuming and challenging in some sections, many researchers have suggested the use of Unmanned Aerial Vehicles (UAVs) for this purpose [38, 39, 40]. Golparvar-Fard *et al.* [34] developed an ML method for detection of the ongoing progress of construction. They utilised an image data set, including progress and no-progress predefined photos and trained a classification model to be used for prediction of new images. Image classification was also used for the detection of construction materials and building a BIM model to support automated progress monitoring [41].

It was argued that employing image synthesis methods can help improve the accuracy of classifiers. Soltani *et al.* [42] demonstrated the efficiency of synthetic images in training vision detectors of construction equipment. Rashidi *et al.* [43] compared three different classification models for material detection and managed to automatically detect concrete, oriented strand board and red brick from the still images. Kim and Kim [44] used the histogram of oriented gradient visual descriptor for training a set of synthetically created images to help facilitate site equipment detection.

2.3. Virtual and Augmented Reality for Construction Project Monitoring

Initially used in the gaming industry, VR applications have long since entered the architecture and construction industries, allowing better visualisation and simulation of various scenarios [45]. Kim and Kano [46] advocated the superiority of related VR images over the ordinary photographs taken from the construction sites for progress monitoring. It has been argued that VR images can provide a realistic location and condition of structure elements comparable to 3D CAD models [47]. VR technologies have widely been used for design and construction prototyping by modelling and visualising different activates, in order to identify potential risks and optimise construction process [48], also to help effectively manage design alterations and improve communications with clients[49]. Retik *et al.* [50] developed a hybrid VR interface integrated with telepresence and video communication systems allowing remote construction monitoring.

AR has been getting more attention in the construction industry because of its ability to superimpose virtual objects on real world scenes. Several research works have used AR for providing more accurate interactive site visualisation [51], construction worksite planning [52], underground infrastructure planning and maintenance [53], comparison of as-built and as-planned images on construction sites [54], and visualisation of equipment operation and tasks [55]. The effectiveness of using AR tools in supporting decision-making and conveying the sophisticated knowledge to the parties engaged in construction has also been overtly advocated [54].

Kopsida and Brilakis [56]proposed a method to enhance inspection and progress monitoring for interior activities, using HoloLens. The application projected 3D as-planned model on the real-world scene, for identifying inconsistencies with actual construction. Ratajczak *et al.* [57] integrated AR with BIM and location-based management system, as a mobile application to facilitate the progress monitoring and communication of construction project parties. The app is supported by Tango ready smartphones and displays the 3D BIM model overlaid onto as-built images. It also delivers information on construction tasks and materials technical data.

2.4. Summary of Progress Monitoring Methods

The various methods used for construction progress monitoring is summarised in Table 1, highlighting the limitations and advantages of each.

Table 1: Summary of data acquisition and site visualisation for construction progress monitoring

Type	Method	Advantage	Limitation	Ref
Geospatial techniques	GPS & GIS	Identification of the optimal location for the equipment	Limited to outdoor operation	[16, 17, 18]
	Barcode & QS	Cost-effective, no extra device to read the codes	Time consuming, requires line-of-sight	[58, 19]
	RFID	Do not require line-of-sight, close proximity, individual reading and direct contact	Prone to error in presence of metal and liquids, time-consuming	[21, 22, 20]
	UWB	Reliable signals, low energy, provide 3D position coordinates	High cost and no mini device or daily-necessity-embedded tool	[22, 59]
Imaging techniques	Laser scanning	Automated data collection, high accuracy	Expensive equipment, laborious	[26, 23, 24]
	Photogrammetry	Calculation of completion percentage and measurement the project progress	Sensitivity to lighting conditions, high time complexity	[60, 61]
	Videogrammetry	Low time complexity, detection of moving equipment	Lower accuracy than photogrammetry	[62, 63]
VR/AR techniques	VR from site photos	provide realistic location and condition of structure element, remote construction monitoring	Inability to check with as-planned model	[50, 46, 49]
	AR	Superimpose as-planned on as-built image	No support for remote monitoring, and automated actions, depends on the device technologies	[54, 51, 55]

2.5. Research Gap

Although there have been tools including geospatial and imaging techniques to enhance progress tracking of construction projects, these applications are not yet able to effectively identify the inconsistencies between as-built and as-planned models. Moreover, there is a need for a decision support system for project monitoring to support effective communication among the involved parties. The reviewed literature shows that there has been a great success in employing VR and AR applications for supporting decision-making at design stages and handling complicated tasks in facilities management. However, this great potential has not been fully utilised in the construction phase for enhancing project progress monitoring and reporting, mainly due to specific conditions of construction sites which make them different from any other environment. It was mentioned that previous studies have been successful in integrating AR and BIM models to provide a tool for construction progress inspection. However, these tools heavily rely on the accuracy of utilised AR tools, and thereby prone to error. On the other hand, the use of such solutions for proper communications among construction parties requires the presence of contributors on construction sites. This issue is one of the main raising challenges in construction management. The use of VR to tackle this problem has been very limited as it requires multidisciplinary cooperation to realise the interoperability of it with 3D information modelling and advanced AI methods.

The framework and prototype presented in this paper aim at integrating ML, artificial intelligence, image processing, VR and BIM technologies aligned with gamification approaches in order to address this particular research gap. Unlike previous attempts that utilised ML and BIM directly as a means for identifying work progress or diagnosing particular problems, this framework benefited from those technologies for the preparation of an interactive virtual environment to be manipulated by a game engine. Therefore, this tool allows effective utilisation of these technologies in order to support selective examination of various building characteristics at different times and by different people, hence making the system more usable for all professionals involved in the project.

3. Project Framework

The core component of the hybrid system described in this article is a game-like VRE, providing integration between ML-based image processing and BIM. As such, the developed system architecture in this study consists of four major elements: 1) image capturing, 2) image processing, 3) BIM authoring, and 4) VR game authoring and object linking. Overall, the study proved the concept that the integration of computer vision, image processing, BIM and VR can facilitate the automatic update of a digital model, storage of the data in a standard file format, and display of project progress information in a structured manner. This paper posits that this can be used as an effective tool for communicating with different parties involved in construction projects and decision-making.

Despite the fact that laser scanning and photogrammetry are the leading methods for collecting special and geometric information, still image photography method was adopted in this study, as an inexpensive and hassle-free method, that has become the industry standard for gathering construction site progress information.

Figure 1 demonstrates the schematic diagram of the proposed system architecture. Interoperability and flawless information exchange were seen as a significant driver in this research; therefore, Autodesk Naviswork (NWD) file format and IFC as standard file format were used throughout this research. First, entities from BIM IFC model links to the unity to create the VR model. From this model, a dataset of synthetic RGB-D images is generated. Then, Neural networks are trained using the created data enriched with real-world depth images to create classifier models.

Construction site images were captured every day using a depth camera, then regenerated by copying the same camera settings and location (localisation) within the BIM model. These images were stored in cloud-based storage along with the BIM model. ML classification technique, Convolutional Neural Network (CNN), was applied to the aforementioned photographs to detect and identify different objects and building components. Image processing was then utilised to remove unwanted objects from the scenes for further clarity.

Recognised and classified elements were linked to the related actionable tasks from the time schedules linked to the BIM model. The extracted details were then overlaid onto the as-planned BIM models for the purpose of comparison.

The superimposed construction status data was transferred to the game engine through scripting or exporting IFC into an Autodesk FilmBoX (FBX) format, then integrated with the virtual environment. This provided the system with a higher level of immersion, improved visualisation and enhanced interaction. The VRE contained both as-planned and as-built models for the purpose of comparison and identification of potential discrepancies.

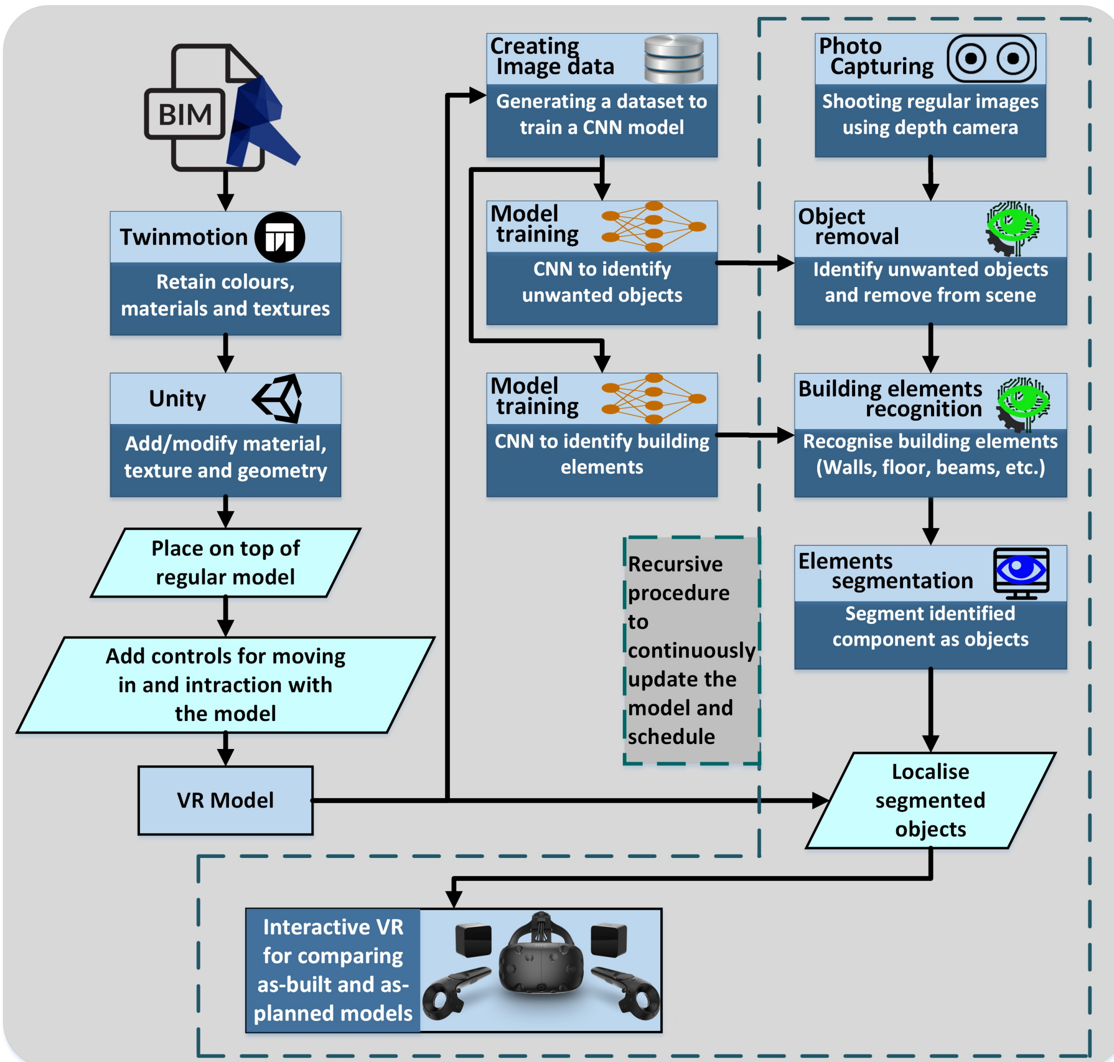


Figure 1: Flowchart of the proposed framework for construction progress monitoring using BIM, VR and image processing technologies.

The proposed techniques will be elaborated in the following sections by providing detail of prototype development.

4. Prototype Development

The prototype of the application incorporating ML-based image processing, BIM and VR were implemented in order to showcase the feasibility and potential of the proposed conceptual framework. The research selected the new leisure and sports complex of the University of Strathclyde located in Glasgow city centre. The structural design of this building was exported to Autodesk Naviswork (NWD) file format, and the architectural model was saved in IFC format which made it suitable for unobstructed data exchange between different BIM authoring software applications, used in the building industry by various parties. Figure 2 (a) and (b) respectively, present architectural and structural BIM models of the mentioned building from the same perspective.

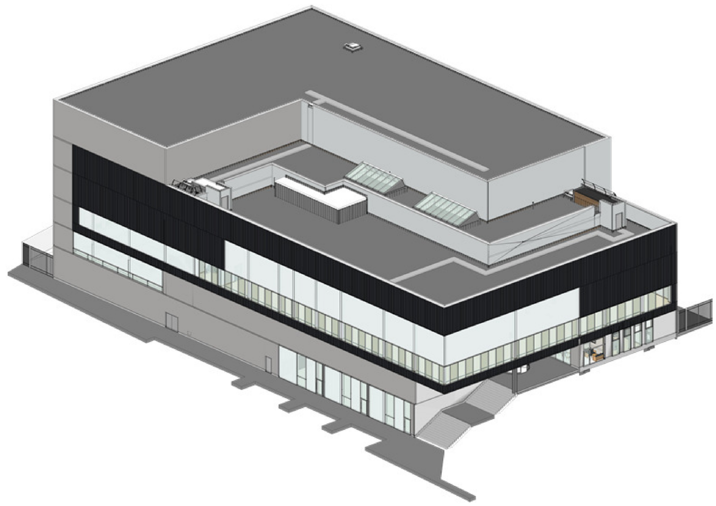
4.1. Image Processing

For the purpose of this prototype, this study developed an optimised approach to ML-based image processing for the automatic detection and recognition of the main constructional and structural elements. In order to support a neater virtual environment, the study also developed a method for removing unwanted objects from the scenes. The developed image processing method starts with a depth camera acquiring multiple overlap RGB-D (colour+depth) images as system inputs for object detection and identification. The primary method we used for object detection and image segmentation is based on ML.

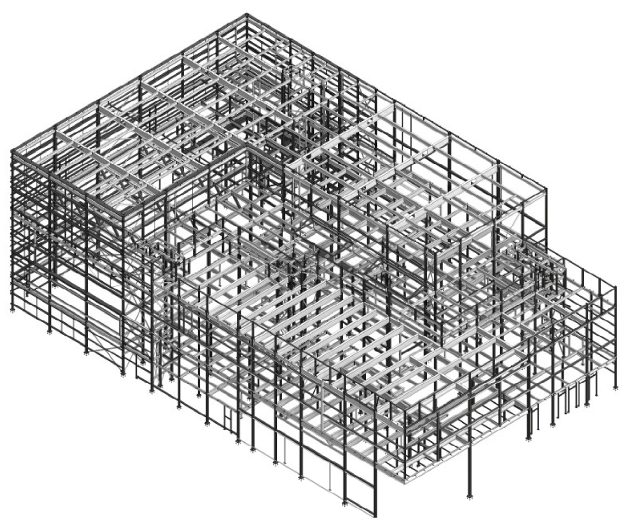
Image segmentation, which partitions a set of pixels into multiple groups, was employed to designate building elements in the image coordinates. This method is widely investigated in many practical applications, such as video surveillance, image retrieval, medical imaging analytics, object detection and location, pattern recognition, etc. [64].

By using a 2D BIM image/video, physical coordinates can be estimated, and point cloud or depth image can be generated via structure from motion (SfM) technique [65, 66]. In the BIM model, a pre-defined 3D data is provided as the prior knowledge, which can supervise the machine to recognise BIM objects via ML algorithms, such as support vector machine (SVM), random forests (RF) and CNN.

A carefully designed dataset is required for training a robust classifier. Image processing for pose estimation [67] and semantic segmentation of the



(a)



(b)

Figure 2: (a) Architectural and (b) structural BIM models.

scene (Song *et al.*, 2017), includes labelling each pixel within the depth image with an object associated category to supervise the ML algorithm and learning the voxel features of the object. In BIM projects, it is a practical approach to use a pre-defined 3D model for simulating depth images in various viewpoints, thus generating as many labelled depth images as possible for building a dataset prior to the commencement of the construction work. The details of generating data for training the network using virtual photogrammetry are presented in the next section.

Taking daily construction activity planning into consideration, a deep neural network seems a proper option, as it can automatically tune the parameters [68, 69], based on the new given data. Therefore, it can effectively learn a new representation of an object along with the progress of a BIM project instead of retraining a new classifier on a daily basis.

Due to the diverse nature of predictions for semantic segmentation and recognition of unwanted objects in the construction scenery, training a single model to perform both tasks accurately is rather difficult. Moreover, the utilised dataset for training a model for image segmentation which includes synthetic images does not include those intended objects, such as human or machinery. As such, this study first applied a separate network for recognition of these objects. This procedure guarantees the precision of both models by utilisation of individual training sets. It is possible to train one model for both purposes, using a comprehensive dataset. However, preparation of such dataset is very laborious as it requires a graphical mixture of environment and objects, both in the form of synthetic and real images. Moreover, obtaining adequate negative data is another hurdle for creating a recognition model. It should be noted that the negative samples are as valuable as positive records in training an accurate model that requires identifying the desired objects precisely as well as rejecting the false detections.

4.1.1. Object removal

Taking photos from a construction site is the first step in creating an as-built model for the purpose of comparison between the current state and the as-planned model. Construction scenes consist of many tools and materials, which are considered as unwanted objects in the construction progress monitoring application. Furthermore, the presence of these objects will lead to faulty detection of construction elements related to the BIM model. Generally, it is not practical to move all these objects while shooting photos. This study proposed an automated two-stage object removal method in order to address this challenge. The first step was using the supervised object recognition technique for identifying unwanted objects, and the second step was to fill the area of the detected object in a visually plausible way.

Many approaches for filling a part of an image or inpainting have been developed by the previous studies. However, as the images used in the study contained depth information, it was posited that a suitable region filling method should be able to estimate the depth filling pixels. This study employed an exemplar-based method developed by [70] as a suitable means to address this requirement. First, these elements were recognised and segmented using the pre-trained CNN, then the boundary of the target region was identified, a patch was chosen to be inpainted, and the source area was queried to find the best-matching spot via an appropriate error metric. This study noted that the main advantage of this method over the other inpainting methods was its ability to propagate the texture into the target region.

Gupta *et al.* [71] trained a large CNN on RGB-D images to recognise objects and applied a semantic segmentation to infer object masks. For the purpose of demonstration, this study trained CNN to detect humans in the construction scene. For training data, the study selected 3200 positive images [72] and 7500 negative images [72]. Figure 3 shows the detail of the method for recognition of objects using RGB-D images and presents how RGB-D images make it possible to calculate the depth and average gradients. This detail is then combined with a fast edge detection approach to generate enhanced outlines. The contours are used to create 2.5D region candidates through processing characteristics on the depth and colour image. The depth-map is encoded with various channels at each pixel, namely horizontal disparity, height over ground, and the angle the pixel's local surface. Then the CNNs which are trained on RGB-d images detect intended objects in 2.5D region candidates. Each CNN begins with a set of region proposals, and calculate features on them. The box proposals are then classified using a linear support vector machine.

Figure 4 shows the original image and with a depth-map, taken from the construction site. Figure 5 presents the output of applying object recognition with the trained CNN.

When the unwanted object is recognised, the filtered image is passed to the object removal procedure to eliminate it from the scene and fill the gap in both RGB and depth space. The outcome of applying this algorithm in Figure 4 is demonstrated in Figure 6. The depth-map of the generated image is illustrated in Figure 6 (b).

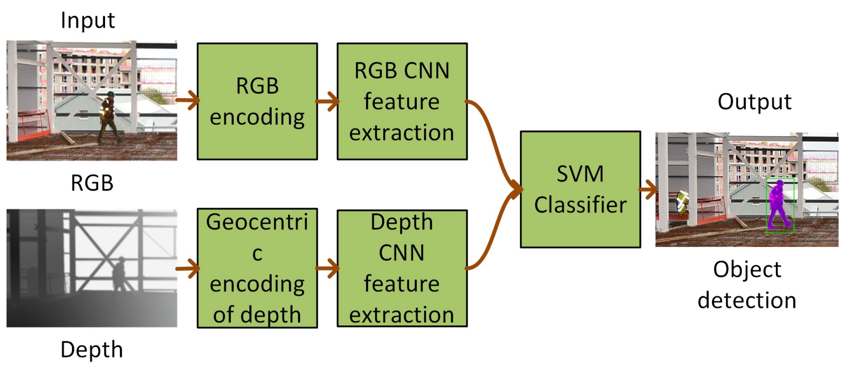
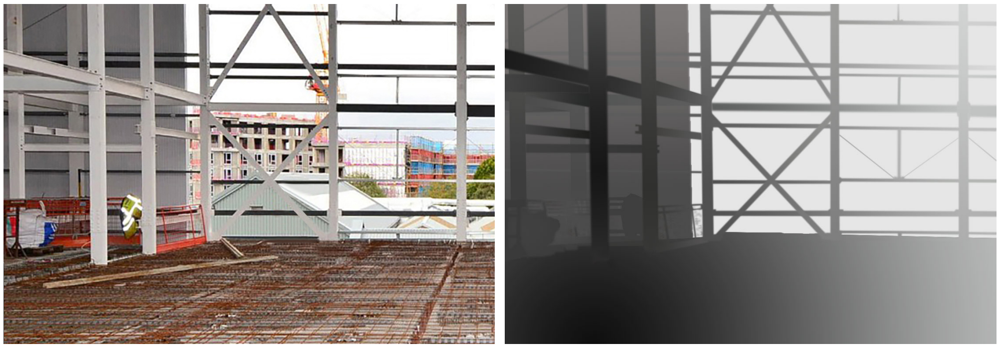


Figure 3: Object recognition using colour and depth images of a scene.

(a) (b)

Figure 4: (a) A photo taken from the construction site and (b) it's depth-map.

Figure 5: Creating an object mask with human recognition using CNN.



(a) (b)

Figure 6: (a) A photo taken from the construction site and (b) it's depth-map.

4.1.2. Building elements identification

The next step in the preparation of the taken depth images for being used in the VRE was to apply a semantic segmentation method to identify the various building elements. In this step, the pixels are directly labelled, to create the segments using a trained CNN. The training data for this network is generated from the BIM model, adding depth information and pixel labels. The use of synthetic images with the aim of semantic segmentation has been widely reported [73, 74, 75], however, due to perfection of those images, the networks may fail to learn all characterisation of the noisy real photos. Moreover, as mentioned before, there was already a need for training another network for the detection of unwanted objects.

For building scene segmentation, this study adopted the FuseNet algorithm [76], as a fully fusion-based CNN developed for semantic segmentation of RGB-D images. After training the ML model with the generated dataset, the model was applied to the image resulted from object removal. Figure 7 demonstrates the architecture of the network for semantic segmentation. The network contains encoders for obtaining features from RGB and depth images and a decoder which maps the feature into the original input resolution. Afterwards, the elements from the depth encoders are fused into the feature-maps of the RGB part.

Figure 8 presents the outcome of the segmentation with the coloured areas indicating different building elements.

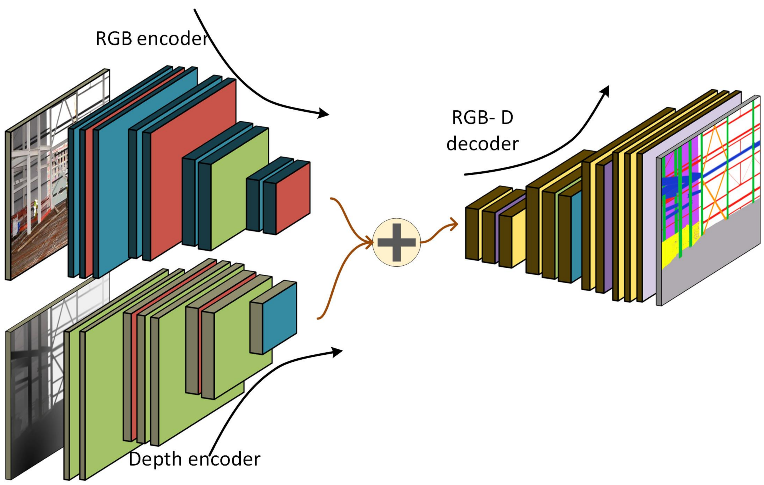


Figure 7: Semantic segmentation by incorporating depth into colour image.

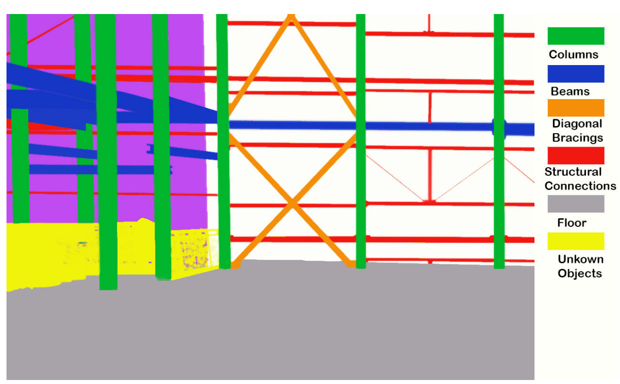


Figure 8: Output of semantic segmentation for identifying structural building parts.

4.2. Linking the IFC Model and the Unity Game Engine

This study used the Unity game engine [77] as the main platform of integration of various technologies, i.e. BIM, ML-based image processing and VR. For ease of communication, in the remaining pars of this paper Unity game engine will be referred to as Unity. This study adopted a previously developed method of linking a digital model [14] to support the integration of the information contained in the BIM model. The method was established based on the use of IFC file format and the incorporation of programming with C#.

Entities from IFC models were linked to Unity through procedurally generated regular expressions, where there are no families. They were, however, combined with spatial localisation, where families are present. The primary challenge in this process was due to the geometric representation characteristics and Unity equivalents. This required a supporting planar geometry intermediary library linking Unity's plain ignorant, mixed Cartesian and barycentric coordinate systems with coplanar, relative Cartesian coordinate systems. To tackle this issue, the system mixed ambivalent interactions with geometry and took advantage of the game engine, in order to optimise interactions with the environment.

Naming conventions between IFC model, FBX, and Unity environment models are found to differ unpredictably depending on the applied export-to-import processes and the source application. Several characteristics of the naming convention in Unity, typically swapped delimiters, replaced, removed and amended characters, and varied in case sensitivity retention which prevented direct one-to-one linking or effective character-by-character comparison. This issue was resolved using procedurally generated regular expressions and iterative searching over the entity recordset. The model was extended to include a top-level record array with records either without parents, omitting representation layers, or entities presumed to be physical entities which significantly reduced the processing search set. This subsequently reduced processing time.

However, families presented two additional challenges for linking. This was partially due to categorical naming returning a set of candidates, whereas familyless comparisons returned single associations. It was also because of the need to naming family-named entities, hierarchically, which required stepped iteration through nested game objects for full name generation. Nevertheless, due to failing to identify a non-spatial attribute to filter candidates, identification of the appropriate representation relied on localisation. This task was achieved by using bounding box centre points, which were generated by Unity during mesh construction and lazily evaluated by the IFC Library.

Figure 9 illustrates the procedure of integrating BIM to VR and linking segmented photos and the flow of information between each step.

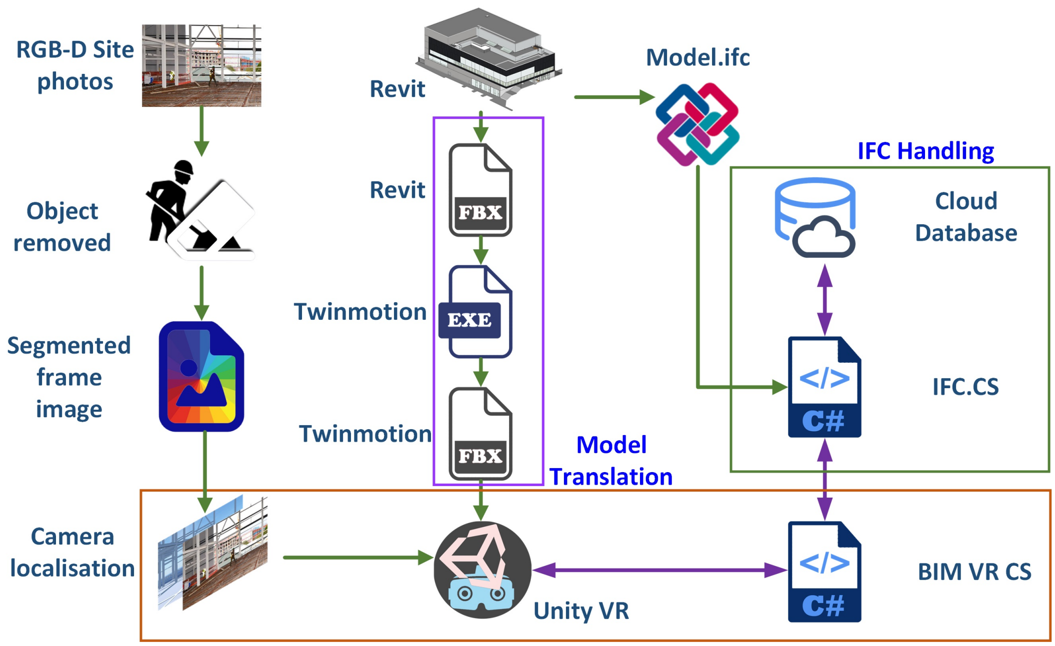


Figure 9: Flowchart of creating VRE from BIM model and overlying real world images.

4.3. Virtual photogrammetry for NN training

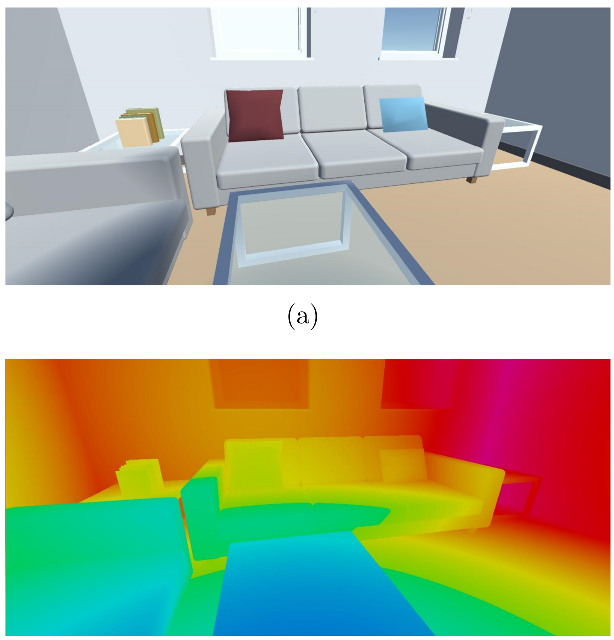
Point cloud generation in the real world has several notable forms including LiDAR, laser scanning and Structure-from-Motion (SfM). Each method produces a vector map of points representing the physical elements and pixel information from the device's point of capture. Virtual worlds in terms of physical representations discrete interval scale and rendering definition, are not meaningfully different from the real world. This enables the application of many traditional and contemporary photogrammetry methods with varying levels of algorithmic complexity. However, while many would be transferable to the virtual world, the real world techniques can rely on imperfections in the collected images. For example, in a large zone in a virtual world rendered with acrylic paint, there may be few if any identifying features, instead of a featureless image with constant colour pixels. Although this is rarely a case in virtual models, a method needs to work with everything or nothing. This reduces the potential for SfM. Similarly, implementations of real world tools in a virtual world are often significantly different from how they function in reality. LiDAR, for example, uses omnidirectional pulses. If this were implemented literally in the virtual world, a countable but impractical number of ray casts would be required which would bring the engine to a halt for a long time.

Instead, at point of triggering the function in the virtual world, the triggering objects that reside and interrogate a physical object database would step out of the world. For instance, in the case of a game with human agents, rather than raycast in every direction, the script would locate all players in the database then separate the (x,y) and (z) dimensions. Using the (x,y) boundaries, a simple rectangle would be created around the agent, and an intersection calculation would be carried out for each line until a horizontal intersection occurs; otherwise, the play is ignored. Assuming it is identified, in the next stage, a vertical intersect assertion would be made. The simplest solution at this point may ray cast only within the boundaries of the box. A more complex solution may use relative coordinate systems and spatial caching to reduce further the number of cases required. For example, for any given intersect the boundary box method may be applied to infer points which do not merit raycasting.

In order to avoid a complicated algorithm based on existing techniques used in the real world, this study facilitated virtual photogrammetry using a simple application of the game engine's physics component and its camera class functionality. Using the Camera class's camera point to virtual world coordinate translation function, each pixel's relative location in the virtual world is identified. A raycast is sent from that location, matching the base orientation of the camera then attempts to intersect with any object with a MeshCollider. Upon successfully hitting an object, a struct is created containing the 2D and 3D coordinate system locations, an extension to the default Mesh class which accommodates binding of external data, such as IFC and the triangle on the mesh which was hit. The latter implicitly linking a Barycentric coordinate system to the struct. As mentioned in the planar geometry system, triangular meshes are translated into planar surface sets. This set is already bound with the mesh, and each plane constituents mesh triangles. Each point struct is added to the point cloud dictionary, which may then be exported for producing composite images, or helping scalable discrete reconstruction.

The library is susceptible to two issues common with other forms of photogrammetry, i.e. nonuniform intervals and point density. However, in contrast to the real world, the process is reproducible and does not require an additional journey through the model. At any point, where the cloud lacks integrity, a camera can be spawned to collect additional data with partially controlled precision. Partial is used here as a caveat for the inherent limitations of casting a finite number of times on varying distance surfaces.

Figures 10 (a) and (b) show the camera view and spatial linking between image pixels and mesh continuous spatial information from the localised camera and BIM model. Every pixel represents an object containing 3D coordinates, and existing and mapped colouring. .



(a)

(b)

Figure 10: (a) Camera and (b) spatial distance heat mapped (depth) views.

Figure 11 demonstrates applying discrete spatial constraints on virtual environment translation. Spatial data was decoupled into continuous, discrete and photogrammetric vectors, representing the real position of the object in the virtual environment, interval equivalent of that object, and its position as it appeared in the primitive reconstruction.

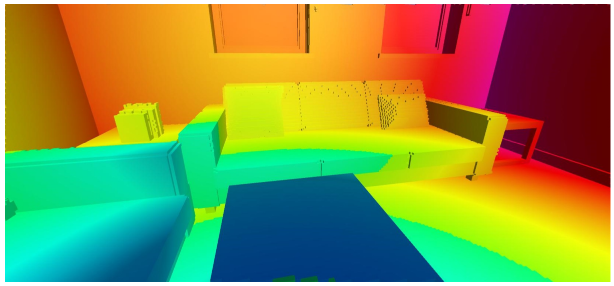


Figure 11: Discrete distance Heat-mapped view.

Synthetic image data is collected primarily using the virtual photogrammetry methods demonstrated in images 10 and 11, and procedurally generating random camera locations in spaces or near the objects of interest. Using the floor perimeter, the camera location boundaries can be obtained, and the number of floor elements can be used as a proxy. Cameras are first spawned around human height, enabling initial data collection regarding the current space, primarily making ceiling height identification and partially identifying zone boundaries. Once these have been identified, a camera can store an image and use the learned spatial boundaries to choose a new camera location within the space. The virtual photogrammetry and camera movement in the space can then be iteratively applied to the partial mapping, ceiling height and farthest surface distance to constrain camera angle.

The virtual photogrammetry is then used to bind pixels to their constituent entities in the virtual world, and if bound, to update IFC representations. Their material, entity type and where relevant, the families are extracted from the IFC schema. This is bound to the raw image data included UV and ARGB. Pixels are grouped by the entity that they are associated with. Entities with pixel adjacencies can have surface poly-boundary points check for coplanarity with surfaces on the other entity. The relationships are tested in each imaging set, ensuring most are encaptured. As shown in Figure 12, entities can be turned off such that the surrounding hidden (or partially hidden) entity surfaces can be mapped. If necessary, this could also be used to identify which entities are present in the space, but occluded.

Data is then split into its constituent sets and inserted into a database or tracked in an image relations file. The former interrogated via SQL and the latter parsed into the PixelRelation class of the VP library. Splicing in a localised real-world image now has a truth network for training real-world

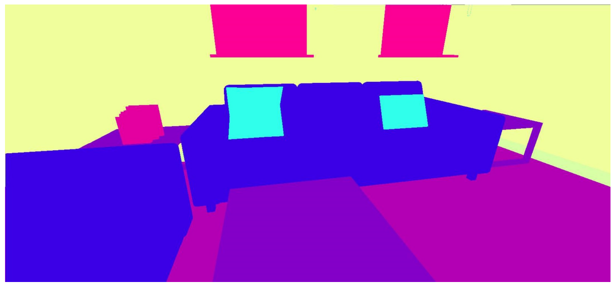


Figure 12: Category-mapped scene view.

classifiers and segmentation models, where the data in the virtual world is appropriate mostly for producing inference networks. The significance of the former is that project like ImageNet are massive human endeavours. People manually segment images and classify objects. Since the virtual world has already segmented virtual representations, these can be used to estimate confidence in classifications. Using depth from the real-world and virtual-world relative scale can be used to refine the expectations. If something seems right, but not linked between both, other profiles from virtual images can be scaled and compared. If an object is still not identified, a case for propagating a specialist classifier for that type of object exists. The family or entity's related images can then be used to train a model looking for those exclusively.

4.4. Integration with the Game Engine

Autodesk Revit viewed and amended the original BIM model of the Strathclyde Sports Centre. The digital model was imported into the Unity game engine to allow for the development of VR simulation. The Unity game creation engine was the primary platform for creating the application, using the C# as the main programming language. The IFC BIM model was exported in the FBX file through a TwinMotion plug-in for Revit. The plug-in was incorporated in the workflow, since the export of the FBX file through Revit would result in loss of the textures defining the materials during the transition to Unity. This method allowed flawless export of the BIM model with the textures assigned. The use of TwinMotion was proved to be very beneficial and contributed to achieving a high-quality asset in Unity. Specific optimisations were made to the BIM model to correct the existing errors in the geometry. The size of the model caused a sort of issues since the entire model and all the objects were rendered at the same time.

The HTC Vive was the primary VR headset used for this work. The Vive comes equipped with two controllers, and two wall or tripod mountable scanners that allow the user to move around extensively within their personal real world space. A Windows gaming laptop with an Intel i7 processor and NVidia GTX 1070 graphics card was the machine used to create the application.

4.5. Overlaying the Segmented Images

When the virtual environment was built in the Unity engine, the final step to complete the monitoring tool was to overlay as-built images over the as-planned model. The elaboration of the method as a set of rules was as follows:

- Where no high-precision camera location information was known, but the entities of interest were identified, they may have been highlighted entirely via the *Colour* property of the primary material. Changing this property would highlight the desired objects, so that when the user brought them into view, they would have been able to identify whether or not they were of interest.

- Where camera location was known, but orientation and field of view weren't, faces could be identified via generation of a FaceSet through the UnityIFC library, which could create the planes to define the entity. In this case, the script needed to effectively raycast to the centre of the closest triangle on the largest two faces. Once the face was identified, the barycentric coordinate systems needed to be used to determine which triangle is associated with that specific raycast. The triangle point list direction which informed the user as to whether the triangle is visible or not rendered. If the triangle is rendered, then the face is that the user should see; otherwise, the other largest face is that of interest. The identified face's primary material colour would then need to be changed to the highlighting colour.

- Where all the above (including the field of view) were known, the system would provide an indication of entities which were not entirely in place. That is not to say the system has definitely identified erroneously positioned columns. However, they might have been highlighted, where a re-measure with a laser distance meter (as Leica Disto) would have been worthwhile. There were two options for achieving this goal: 1) by using superimposing the segmented image over an n-dimensional image created with the library designed for this project, or 2) by applying a similar process for direct raycasting from the superimposed segments. In the case of the former, the screen or secondary camera resolution needed to be adjusted to match the metadata on the picture; or if not possible, up to a proportional scale. It should, however, be noted that in this case, additional control for pixel scale would have been required as well as including buffering the AND failure area to accommodate the larger pixels. Pixels from each then needed to be first compared with an AND operation to identify where expected pixels overlie with the segmented sections from the image processor.

- Where the AND failed, there was a potential misalignment of the entities, and each should have been highlighted appropriately using two distinctive colours to demonstrate where the segment and virtual face were out with the boundary of the other.

Figure 13 illustrates the overall proposed procedure for overlaying the segmented images on the BIM model in the VRE. Figure 14 shows the result of superimposing the segmented image over BIM model in VR environment, where the columns are selected for the investigation.

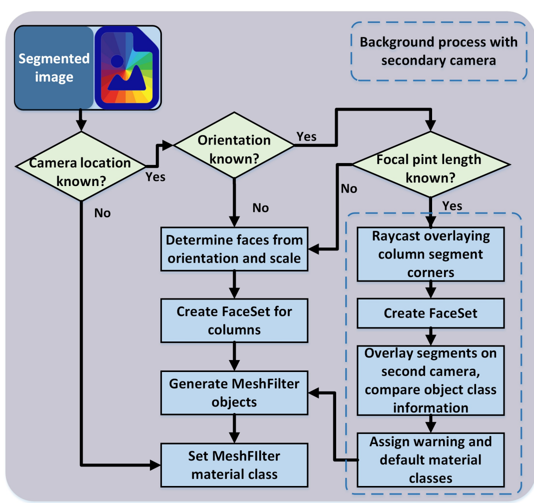


Figure 13: Procedure flowchart of overlaying segmented images over BIM model in VRE.

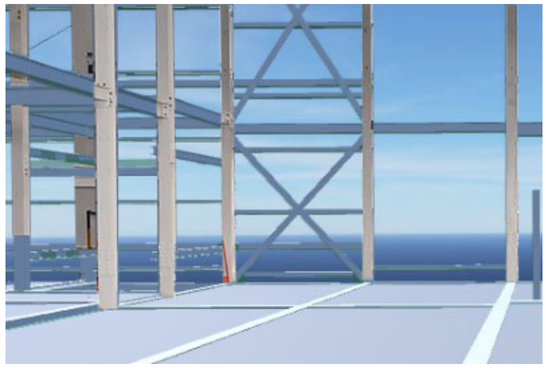


Figure 14: Superimposing the columns in the VR environment.

Camera localisation has several meaningful opportunities depending on the information of entities captured within the view and segmentation. The methods appropriate for this paper primarily rely on segmentation and BIM localisation. Depth segmentation is mostly excluded here, but confidence in its results can surpass the random sampling type method described. The following describes an idealised (two-sample) instance, however, the principle of RANSAC could be applied in conjunction with it to refine and confirm.

IPS has been reported to be accurate within 300mm which is a significant improvement on GPS, reported to be accurate within 5,000mm. In recent years, research has refined the use of fiducial markers, demonstrating accuracy within 80mm. In each case, the reported accuracy is without the benefit of verification with segmentation or virtual photogrammetry from BIM models. With the assumption that a floor is segmented, choosing three points on the surface, with two creating a line perpendicular to the camera, forms an isosceles or equilateral triangular hyperplane. Splitting this into a triangle and trapezoid and injecting the z dimension from segmentation will provide enough data to estimate the shear coefficient. Repeating this process on a second perpendicular triangle will only be consistent, if it falls within a reasonable tolerance, and if the surface is horizontal. Otherwise, it should be possible to estimate the second shear coefficient and then rearrange shear functions to determine perpendicular horizontal axis orientation. With the orientation of the surface and shear coefficient, the cameras' height and orientation can be determined. The results may be confirmed by repeating the process on the virtual planes associated with the segmented surfaces by using the pixel reference to identify the surface.

With or without the previous calibration for height and orientation, classified segmentation may aid location inference. Using segmented entities and IPS virtual photogrammetry can be used to search for similar or expected profiles. Considering a beam, for example, the orientation can be determined from another hyperplane, this time between the floor and vertical centre point of the beam, shear coefficient and taking the ratio of the full height and the second height of another point along the floor/beam, not the entire length. By repeating the process, the results may be refined. The final stage for this method is subsampling segmented and virtual properties and applying RANSAC to make educated guess from what appears to be in both from segmentation in both virtual and real worlds.

Another option is using fiducial or natural markers, specifically for localisation of the camera. While natural markers are not necessarily present, anyone taking pictures may throw temporary markers to the floor which, will enable the previous method in conjunction with IPS. Segmentation aided localisation is a matter of confidence in what has been identified. Any markers known to be on the floor will facilitate the previous method. In either case, if a floor is known, IPS accuracy is enough to start using surface-to-surface edge detection connected, noncoplanar surfaces. Localisation is not an exact science and under normal circumstances is not guaranteed to work. However, with BIM overlaying, the virtual test frame can be moved and the VP can be regenerated. It is not great to say educated guessing can solve the problem, however, it doesn't matter whether the process is achieved by a single process or it takes a thousand. Localisation doesn't need to happen in seconds or minutes, it just has to fall within an acceptable tolerance and not require human input.

5. Discussions

Technology-Supported Integration of Real and Virtual Worlds: Bridging the gap between real and virtual-worlds is no longer a matter of technological barriers but rather a resource management issue. The tools which are necessary to link, translate and fuse mixed reality data exist in isolation. It is down to development teams to figure out how they may be combined to produce a practical utility under the constraints of funding and team capacity. Some techniques, such as Drone LiDAR scanning have long since been established for producing a discrete spatial mapping of the real world, that can be translated into virtual components. To a lesser extent, some tools such as Tango-enabled mobiles can produce and translate point-clouds to coloured 3D meshes in near real-time. Pour Rahimian *et al.* [14] demonstrated translation between vector and discrete virtual environments and linking Cartesian and Barycentric coordinate systems. In this study, bidirectional linking between discrete and vector worlds were extended to incorporate raster representations, ultimately developing a virtual photogrammetry tool for mixed reality.

ML-Assisted Image Processing: Python's SciKit-learn contains many flexible deep learning utilities designed for image classification, object identification and semantic segmentation. Of course, ML and deep learning tools aren't without flaw, but through progressive interactive training with reinforcement, transfer learning and input homogenisation, they can be convinced to perform well beyond expectations. Linking virtual and real worlds, however, requires testing, tweaking and reinforcement, which under normal circumstances, would require significant team involvement which cannot always be consumed in parallel and must be carefully managed.

Paradigm Shift in Project Monitoring: The process of developing and proving solutions for research objectives similar to this study do not need to rely on the real world initially. Through abstracting the process, there is no reason why individual components from virtual and real world utilities or data sources cannot be interchangeable. For example, virtual photogrammetry module for the Unity game engine, which was developed in this study can differ from real photogrammetry only by imperfections in the surfaces they may scan. Although virtual models can produce ideal surface maps, there is a little challenge in introducing imperfections. The idea behind converging realities is that tasks which are constrained by serial time, equipment allocation and team capacity are not subject to the same constraints. If a team has one drone in the real world, they can generate data specific to the target building(s) at the rate which the camera is capable of capturing UV and point-cloud data. The properties of the building and environment are immutable, the weather and lighting are situational, and the rendering is not controllable by the pilot.

Expedited Data Collection: In contrast, virtual worlds are constrained only by the amount of processing time and devices that can be afforded to the project. Data collection can be automatic and in parallel, and the environmental and target features are mutable. The virtual world can be spawned randomly with rendering material types, resolution and imperfections unique to a given instance; even the target can be sampled from a repository of buildings. Drones in the virtual world don't need pilots, nor do they have physical representations meaning more than one can collect data during a single collection. Unity can be compiled for Linux machines which enable next-to-nothing cost parallel processing. In short, by the time the real world pilot has travelled to the site and generated data for a single target, the virtual world can produce thousands of data sets from any number of targets with a flexible selection of the characteristics that are otherwise immutable in the real world.

Converging Realities: The aim of attempting to converging realities is to take actions which are applicable to virtual world data applicable to data from the real world by gradually blurring the lines between the two. The real and virtual world renderings are never going to be identical, and therefore, training models on purely virtual data alone would likely be unfruitful. However, between readily available rendering material images and Fast Style Transfer (FST) CNNs creating a progressive interactive training ensemble can be made easier. FSTs, as demonstrated by Engstrom [79], learn the common characteristics of a given training image's artistic style by comparing it with thousands of images with distinct artistic styles including works from historically famous artists. Once a model has been trained, an image can be converted from its raw state to a style-transferred equivalent.

The Virtual Photogrammetry Technique: In this study, rather than attempting to create a complicated algorithm based on existing techniques used in the real world, virtual photogrammetry was facilitated using simple application of the game engine's physics component. Currently constrained by the screen resolution, not a physical constraint, any given camera is sent a request to capture what it sees at the time step it receives the request. At the end of the time, step a call back to capture the spatial and physical object data. This process is by no means perfect and not implicitly transferable to the real world, but it lays the foundation of progressing to a practical tool.

Role of Open Standards, Interoperability and Social Psychology: This research suggested the next disruptive innovation in AEC software will not be in the form of cutting edge design functionality but rather by greater consideration for social psychology's role in effective computer-mediated communication. Pour Rahiumian *et al.* (2019) in contrast, argued that focus should be on interoperability with an emphasis on accommodating immersive virtual environments. This paper proposes an intermediate and demonstrates a collection of tools that partially bridge the gap between these two suggestions. Open standards can facilitate interoperability not only between traditional vector CAD vendors but also packages which are not directly linked to the construction industry. Pour Rahimian *et al.* [14], for example, discussed the potential application of their BIM library to reside outside virtual environments entirely. The library does not presuppose that the interface has any graphical interface to the extent that the model may be interrogated without an interfacing script. This project tackles the opposite side of communication problems in accordance to converging realities approach, in which linking real and virtual worlds has many objective and subjective benefits. Through the ability to mix worlds, the subjective perception of proximity between involved parties can be heightened through inherent increases in media richness while reducing the risk of communication breakdown from initially unverifiable conflicts.

Application, Functionality, Impacts and Contributions of the Study: This paper's application in data science is yet to be established. However, it has the potential to be its most significant contribution to scientific knowledge. The framework and prototype presented in this study formed methodological and technical foundations for a converging realities approach to self-propagating hierarchical deep learning ensembles and creating an AI network which aims to progressively introduce real world images to virtual world datasets, via homogenisation of real and virtual world images. A weakness of that kind of process, however, is its initial attempts to rationalise the real world. This project is well suited to mitigating this traction problem. Linking both real and virtual world enables deterministic assertion of the presence of equipment and construction features through spatial comparison of real and virtual photogrammetry data. This serves to solve two problems. First, where an entity is proven to be present in both worlds but not identified by an image classifier, additional images from either world may be introduced to its training data or can be used to produce a child branch in the network. Second, the information in the virtual world is explicitly linked to the virtual entity and therefore, it can be bound to pixels in a deterministic manner. In short, this library's functionality may reduce the need for difficult segmentation while providing a mechanism for reinforcement and transfer learning.

6. Conclusion

The research presented in this paper addresses the methodological and technical gap in the emerging digital analytical tools of machine learning and computer vision and the advanced visualisation media including BIM, and interactive game-like immersive VR interfaces, in order to leverage automation of construction progress monitoring. To achieve this, the study proposed a framework and developed a proof of concept prototype of a hybrid system which is capable of importing and processing construction site images and integrating them with the nD building information models within a gamelike immersive VRE. It was discussed in lights of the reviewed literature that despite the wide adoption of modern technologies in construction progress monitoring, the use of BIM, as an as-planned constructional model, has not been exploited well, due to the interoperability issues among the relevant technologies.

Therefore, this study responded to the necessity of a platform that can support continuous system update and enable construction managers and clients to effectively compare the building construction (as-built) with as-planned BIM model for the purpose of deficiency detection. The resulting prototype, based around the principles of remote construction project monitoring, took advantage of ML and image processing in removing unwanted objects, recognising and extracting main building characteristics, and overlaying these on the corresponding as-planned nD components. VR technologies, including Unity and HTC Vive, provided a virtual environment to allow better user interaction with these elements.

It was argued in the paper that the proposed platform could help construction companies to follow the progress of their work and diagnose any discrepancy without interrupting the on-site operations. This remote managing tool can also save a great deal of time and provide a more accurate comparison of constructed parts with the as-planned BIM model. On the other hand, this powerful virtual environment can be used for presenting the stage of building development to the clients. The facilitated automatic update of the model makes it possible to have an on-demand schedule method without the need for sophisticated wireless-enabled devices. One of the main contributions of this project was providing a technical exemplar for the integration of ML and image processing approaches with immersive and interactive BIM interfaces, the algorithms and program codes of which can help replicability of these approaches by other scholars.

The methods of object recognition, image processing and overlying real images over as-planned BIM model were suggested for completion and demonstration of the presented prototype. However, different approaches might be adopted in future to achieve better identification and segmentation of the building elements, considering the studied construction site situation and demands. Moreover, advanced positioning systems can be employed for easier localisation of such components through extracting camera location. Furthermore, the prospective photo shooting procedures can be further enhanced by means of contemporary UAVs, which are capable of pre-programmed and autonomous aviation and actions. Hence, the utilisation of advanced small drones can automate image acquisition. This is possible through defining a home point and a route in the virtual environment, then adapting it to the real-world location and finally transferring the flight plan to the drone via a waypoint path. As most advanced UAVs encompass 360-degree sensors, they are able to perform safe indoor manoeuvres. The use of these UAVs can also reduce health and safety risks and allow for capturing the site images, even when there are works in progress.

Acknowledgement

The authors would like to gratefully acknowledge the generous funding received from Advanced Forming Research Centre (AFRC), under the Route to Impacts Scheme (grant numbers: R900014-107 & R900021-101) which enabled the team to conduct this study. This work would also not be feasible without the generous PhD funding for the second author, which was co-funded by the Engineering The Future scheme from University of Strathclyde and the Industry Funded Studentship Agreement with arbnco Ltd (Studentship Agreement Number: S170392-101). The authors are also highly grateful to the considerable support from arbnco Ltd in both funding the third author's MPhil study and supporting his professional development over the last decade. The authors also would like to acknowledge the contributions of Dr Andrew Agapiou, Dr Vladimir Stankovic, Frahad Chamo, Minxiang (Jeremy) Ye, Andrew Graham and Kasia Kozlowska from Strathclyde University, during the initial BIM-VR prototype development stages. Finally, valuable advice received from various managers in AFRC, especially David Grant is highly appreciated.

References

[1] R. A. Stewart, IT enhanced project information management in construction: Pathways to improved performance and strategic competitiveness, Automation in Construction 16 (4) (2007) 511–517. `doi:10.1016/j.autcon.2006.09.001`.

[2] S. J. Kolo, F. P. Rahimian, J. S. Goulding, Offsite manufacturing construction: A big opportunity for housing delivery in Nigeria, Procedia Engineering 85 (2014) 319–327. doi:10.1016/j.proeng.2014.10.557.

[3] F. Pour Rahimian, R. Ibrahim, M. N. Baharudin, Using IT/ICT as a new medium toward implementation of interactive architectural communication cultures, in: Proceedings - International Symposium on Information Technology 2008, ITSim, Vol. 4, IEEE, 2008, pp. 1–11. doi:10.1109/ITSIM.2008.4631984.

[4] E. O. Ibem, S. Laryea, Survey of digital technologies in procurement of construction projects, Automation in Construction 46 (2014) 11–21. doi:10.1016/j.autcon.2014.07.003.

[5] E. J.Adwan, A. Al-Soufi, A Review of ICT Technology In Construction, International Journal of Managing Information Technology 8 (3/4) (2016) 01–21. doi:10.5121/ijmit.2016.8401.

[6] S. Alsafouri, S. K. Ayer, Review of ICT Implementations for Facilitating Information Flow between Virtual Models and Construction Project Sites, Automation in Construction 86 (2018) 176–189. doi:10.1016/j.autcon.2017.10.005.

[7] N. Gu, V. Singh, K. London, BIM Ecosystem: The Coevolution of Products, Processes, and People, in: K. Kensek, D. Noble (Eds.), Building Information Modeling: BIM in Current and Future Practice, no. MAY, John Wiley & Sons, Hoboken, 2014, pp. 197–210. doi:10.1002/9781119174752.ch15.

[8] F. Svalestuen, V. Knotten, O. Lædre, F. Drevland, J. Lohne, Using Building Information Model (Bim) Devices To Improve Information Flow and Collaboration on Construction Sites, Journal of Information Technology in Construction 22 (February) (2017) 204–219.
URL http://www.itcon.org/2017/11, Accessed 7th Nov 2019

[9] M. Sheikhkhoshkar, F. Pour Rahimian, M. H. Kaveh, M. R. Hosseini, D. J. Edwards, Automated planning of concrete joint layouts with 4D-BIM, Automation in Construction 107 (2019) 102943. doi:10.1016/j.autcon.2019.102943.

[10] J. Teizer, Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites, Advanced Engineering Informatics 29 (2) (2015) 225–238. doi:10.1016/j.aei.2015.03.006.

[11] S. Roh, Z. Aziz, F. Peña-Mora, An object-based 3D walk-through model for interior construction progress monitoring, Automation in Construction 20 (1) (2011) 66–75. doi:10.1016/j.autcon.2010.07.003.

[12] K. K. Han, M. Golparvar-Fard, Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs, Automation in Construction 53 (2015) 44–57. arXiv:arXiv:1011.1669v3, doi:10.1016/j.autcon.2015.02.007.

[13] K. K. Han, M. Golparvar-Fard, Potential of big visual data and building information modeling for construction performance analytics: An exploratory study, Automation in Construction 73 (2017) 184–198. doi:10.1016/j.autcon.2016.11.004.

[14] F. Pour Rahimian, V. Chavdarova, S. Oliver, F. Chamo, OpenBIM-Tango integrated virtual showroom for offsite manufactured production of self-build housing, Automation in Construction 102 (2019) 1–16. doi:10.1016/j.autcon.2019.02.009.

[15] T. Omar, M. L. Nehdi, Data acquisition technologies for construction progress tracking, Automation in Construction 70 (2016) 143–155. doi:10.1016/j.autcon.2016.06.016.

[16] J. Song, C. T. Haas, C. H. Caldas, Tracking the Location of Materials on Construction Job Sites, Journal of Construction Engineering and Management 132 (9) (2006) 911–918. doi:10.1061/(ASCE)0733-9364(2006)132:9(911).

[17] V. Ptrucean, I. Armeni, M. Nahangi, J. Yeung, I. Brilakis, C. Haas, State of research in automatic as-built modelling, Advanced Engineering Informatics 29 (2) (2015) 162–171. doi:10.1016/j.aei.2015.01.001.

[18] K. Chen, W. Lu, Y. Peng, S. Rowlinson, G. Q. Huang, Bridging BIM and building: From a literature review to an integrated conceptual framework, International Journal of Project Management 33 (6) (2015) 1405–1416. doi:10.1016/j.ijproman.2015.03.006.

[19] T. M. Lorenzo, B. Bossi, M. Cassano, D. Todaro, BIM and QR Code. A Synergic Application in Construction Site Management, Procedia Engineering 85 (2014) 520–528. doi:10.1016/j.proeng.2014.10.579.

[20] H. Xie, W. Shi, R. R. A. Issa, Using rfid and real-time virtual reality simulation for optimization in steel construction USING RFID AND REAL-TIME VIRTUAL REALITY SIMULATION FOR OPTIMIZATION IN STEEL CONSTRUCTION, Journal of Information Technology 16 (September) (2015) 291–308.
URL http://www.itcon.org/2011/19, Accessed 7th Nov 2019

[21] E. J. Jaselskis, T. El-Misalami, Implementing radio frequency identification in the construction process, Journal of Construction Engineering and Management 129 (6) (2003) 680–688. doi:10.1061/(ASCE)0733-9364(2003)129:6(680).

[22] D. L. Jochen Teizer, M. Sofer, Rapid Automated Monitoring of Construction Site Activities Using Ultra-Wideband, in: 24th International Symposium on Automation and Robotics in Construction (ISARC), no. 2, 24th International Symposium on Automation & Robotics in Construction, Kochi, 2007, pp. 23–28. doi:10.22260/ISARC2007/0008.

[23] R. Lu, I. Brilakis, Digital twinning of existing reinforced concrete bridges from labelled point clusters, Automation in Construction 105 (2019) 102837. doi:10.1016/J.AUTCON.2019.102837.

[24] R. Woodhead, P. Stephenson, D. Morrey, Digital construction: From point solutions to IoT ecosystem, Automation in Construction 93 (2018) 35–46. doi:10.1016/J.AUTCON.2018.05.004.

[25] F. Bosché, M. Ahmed, Y. Turkan, C. T. Haas, R. Haas, The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components, Automation in Construction 49 (2015) 201–213. doi:10.1016/j.autcon.2014.05.014.

[26] Y. Turkan, F. Bosché, C. T. Haas, R. Haas, Automated progress tracking using 4D schedule and 3D sensing technologies, Automation in Construction 22 (2012) 414–421. doi:10.1016/j.autcon.2011.10.003.

[27] S. Ladha, R. Singh, Monitoring construction of a structure (jan 2018). URL https://patents.google.com/patent/US20180012125A1/en, Accessed 7th Nov 2019

[28] Z. Asgari, F. P. Rahimian, Advanced Virtual Reality Applications and Intelligent Agents for Construction Process Optimisation and Defect Prevention, Procedia Engineering 196 (2017) 1130–1137. doi:10.1016/j.proeng.2017.08.070.

[29] S. Seyedzadeh, F. P. Rahimian, I. Glesk, M. Roper, Machine learning for estimation of building energy consumption and performance: a review, Visualization in Engineering 6 (1) (2018) 5. doi:10.1186/s40327-018-0064-7.

[30] J. Seo, S. Han, S. Lee, H. Kim, Computer Vision techniques for construction safety and health monitoring, Advanced Engienering Informatics 29 (2) (2015) 239–251. doi:10.1016/j.aei.2015.02.001.

[31] Y. Ham, K. K. Han, J. J. Lin, M. Golparvar-Fard, Visual monitoring of civil infrastructure systems via camera-equipped Unmanned Aerial Vehicles (UAVs): a review of related works, Visualization in Engineering 4 (1) (2016) 1. doi:10.1186/s40327-015-0029-z.

[32] H. Guo, Y. Yu, M. Skitmore, Visualization technology-based construction safety management: A review, Automation in Construction 73 (2017) 135–144. doi:10.1016/j.autcon.2016.10.004.

[33] M. Golparvar-Fard, F. Peña-Mora, C. A. Arboleda, S. Lee, Visualization of Construction Progress Monitoring with 4D Simulation Model Overlaid on Time-Lapsed Photographs, Journal of Computing in Civil Engineering 23 (6) (2009) 391–404. doi:10.1061/(ASCE)0887-3801(2009)23:6(391).

[34] M. Golparvar-Fard, F. Peña-Mora, S. Savarese, Automated Progress Monitoring Using Unordered Daily Construction Photographs and IFC-Based Building Information Models, Journal of Computing in Civil Engineering 29 (1) (2015) 04014025. doi:10.1061/(ASCE)CP.1943-5487.0000205.

[35] A. Braun, S. Tuttas, A. Borrmann, U. Stilla, A concept for automated construction progress monitoring using BIM-based geometric constraints

and photogrammetric point clouds, Journal of Information Technology in Construction 20 (5) (2015) 68–79.
URL http://www.itcon.org/2015/5, Accessed 7th Nov 2019

[36] C. Kim, B. Kim, H. Kim, 4D CAD model updating using image processing-based construction progress monitoring, Automation in Construction 35 (2013) 44–52. doi:10.1016/j.autcon.2013.03.005.

[37] Y. Wu, H. Kim, C. Kim, S. H. Han, Object Recognition in Construction-Site Images Using 3D CAD-Based Filtering, Journal of Computing in Civil Engineering 24 (1) (2010) 56–64. doi:10.1061/(ASCE)0887-3801(2010)24:1(56).

[38] J. J. Lin, K. K. Han, M. Golparvar-Fard, A Framework for Model-Driven Acquisition and Analytics of Visual Data Using UAVs for Automated Construction Progress Monitoring (ASCE), in: Computing in Civil Engineering, 2015, pp. 156–164. doi:10.1061/9780784479247.020.

[39] S. Kluckner, J. A. Birchbauer, C. Windisch, C. Hoppe, A. Irschara, A. Wendel, S. Zollmann, G. Reitmayr, H. Bischof, AVSS 2011 demo session: Construction site monitoring from highly-overlapping MAV images, in: 2011 8th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2011, IEEE, 2011, pp. 531–532. doi:10.1109/AVSS.2011.6027402.

[40] K. K. Han, J. J. Lin, M. Golparvar-Fard, A Formalism for Utilization of Autonomous Vision- Based Systems and Integrated Project Models for Construction Progress Monitoring, in: A Formalism for Utilization of Autonomous Vision- Based Systems and Integrated Project Models for Construction Progress Monitoring, 2015, pp. 128–131.
URL https://core.ac.uk/download/pdf/38930727.pdf#page=129, Accessed 7th Nov 2019

[41] A. Dimitrov, M. Golparvar-Fard, Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections, Advanced Engineering Informatics 28 (1) (2014) 37–49. doi:10.1016/j.aei.2013.11.002.

[42] M. M. Soltani, Z. Zhu, A. Hammad, Automated annotation for visual recognition of construction resources using synthetic images, Automation in Construction 62 (2016) 14–23. doi:10.1016/j.autcon.2015.10.002.

[43] A. Rashidi, M. H. Sigari, M. Maghiar, D. Citrin, An analogy between various machine-learning techniques for detecting construction materials in digital images, KSCE Journal of Civil Engineering 20 (4) (2016) 1178–1188. doi:10.1007/s12205-015-0726-0.

[44] H. Kim, H. Kim, 3D reconstruction of a concrete mixer truck for training object detectors, Automation in Construction 88 (2018) 23–30. doi:10.1016/j.autcon.2017.12.034.

[45] J. Whyte, Virtual Reality and the Built Environment, Routledge, 2002. doi:10.4324/9780080520667.

[46] H. Kim, N. Kano, Comparison of construction photograph and VR image in construction progress, Automation in Construction 17 (2) (2008) 137–143. doi:10.1016/j.autcon.2006.12.005.

[47] F. Pour Rahimian, T. Arciszewski, J. S. Goulding, Successful education for AEC professionals: case study of applying immersive game-like virtual reality interfaces, Visualization in Engineering 2 (1) (2014) 4. doi:10.1186/2213-7459-2-4.

[48] T. Huang, C. W. Kong, H. L. Guo, A. Baldwin, H. Li, A virtual prototyping system for simulating construction processes, Automation in Construction 16 (5) (2007) 1–21. doi:10.1016/j.autcon.2006.09.007.

[49] H. Li, T. Huang, C. W. Kong, H. L. Guo, A. Baldwin, N. Chan, J. Wong, Integrating design and construction through virtual prototyping, Automation in Construction 17 (8) (2008) 915–922. doi:10.1016/j.autcon.2008.02.016.

[50] A. Retik, G. Mair, R. Fryer, D. McGregor, Integrating virtual reality and telepresence to remotely monitor construction sites: A ViRTUE project, in: I. Smith (Ed.), Artificial Intelligence in Structural Engineering. Lecture Notes in Computer Science, Springer, 2006, pp. 459–463. doi:10.1007/bfb0030474.

[51] A. H. Behzadan, V. R. Kamat, Georeferenced Registration of Construction Graphics in Mobile Outdoor Augmented Reality, Journal of Computing in Civil Engineering 21 (4) (2007) 247–258. doi:10.1061/(ASCE)0887-3801(2007)21:4(247).

[52] X. Wang, Using augmented reality to plan virtual construction worksite, International Journal of Advanced Robotic Systems 4 (4) (2007) 501–512. doi:10.5772/5677.

[53] G. Schall, E. Mendez, E. Kruijff, E. Veas, S. Junghanns, B. Reitinger, D. Schmalstieg, Handheld Augmented Reality for underground infrastructure visualization, Personal and Ubiquitous Computing 13 (4) (2009) 281–291. doi:10.1007/s00779-008-0204-5.

[54] M. Golparvar-Fard, F. Peña-Mora, C. A. Arboleda, S. Lee, Visualization of Construction Progress Monitoring with 4D Simulation Model Overlaid on Time-Lapsed Photographs, in: Journal of Computing in Civil Engineering, Vol. 23, 2009, pp. 391–404. doi:10.1061/(ASCE)0887-3801(2009)23:6(391).

[55] A. Hammad, Distributed augmented reality for visualising collaborative construction tasks, Mixed Reality In Architecture, Design And Construction 23 (December) (2009) 171–183. doi:10.1007/978-1-4020-9088-2_11.

[56] M. Kopsida, I. Brilakis, BIM registration methods for mobile augmented reality-based inspection, in: eWork and eBusiness in Architecture, Engineering and Construction - Proceedings of the 11th European Conference on Product and Process Modelling, ECPPM 2016, CRC Press, 2016, pp. 201–208.

[57] J. Ratajczak, A. Schweigkofler, M. Riedl, D. T. Matt, Augmented Reality Combined with Location-Based Management System to Improve the Construction Process, Quality Control and Information Flow, in: Advances in Informatics and Computing in Civil and Construction Engineering, Springer, 2019, pp. 289–296. doi:10.1007/978-3-030-00220-6.
URL http://link.springer.com/10.1007/978-3-030-00220-6, Accessed 7th Nov 2019

[58] H. P. Tserng, R. J. Dzeng, Y. C. Lin, S. T. Lin, Mobile construction supply chain management using PDA and Bar Codes, Computer-Aided Civil and Infrastructure Engineering 20 (4) (2005) 242–264. doi:10.1111/j.1467-8667.2005.00391.

[59] Y. K. Cho, J. H. Youn, D. Martinez, Error modeling for an untethered ultra-wideband system for construction indoor asset tracking, Automation in Construction 19 (1) (2010) 43–54. doi:10.1016/j.autcon.2009.08.001.

[60] Z. A. Memon, M. Z. Muhd, M. Mustaffar, An automatic project progress monintoring model by integrating auto CAD and digital photos, in: Proceedings of the 2005 ASCE International Conference on Computing in Civil Engineering, 2005, pp. 1605–1617.

[61] F. Dai, M. Lu, Assessing the accuracy of applying photogrammetry to take geometric measurements on building products, Journal of Construction Engineering and Management 136 (2) (2010) 242–250. doi:10.1061/(ASCE)CO.1943-7862.0000114.

[62] I. Brilakis, H. Fathi, A. Rashidi, Progressive 3D reconstruction of infrastructure with videogrammetry, Automation in Construction 20 (7) (2011) 884–895. doi:10.1016/j.autcon.2011.03.005.

[63] F. Dai, A. Rashidi, I. Brilakis, P. Vela, Comparison of image-based and time-of-flight-based technologies for three-dimensional reconstruction of infrastructure, Journal of Construction Engineering and Management 139 (1) (2013) 69–79. doi:10.1061/(ASCE)CO.1943-7862.0000565.

[64] G. Dong, M. Xie, Color clustering and learning for image segmentation based on neural networks, IEEE Transactions on Neural Networks 16 (4) (2005) 925–936. doi:10.1109/TNN.2005.849822.

[65] J. L. Schönberger, E. Zheng, M. Pollefeys, J.-M. Frahm, Pixelwise View Selection for Unstructured Multi-View Stereo, in: European Conference on Computer Vision (ECCV), Amsterdam, 2016. doi:10.1007/978-3-319-46487-9_31.

[66] J. L. Schonberger, J.-M. Frahm, Structure-From-Motion Revisited, in: The IEEE Conference on Computer Vision and Pattern Recognition

(CVPR), IEEE, Las Vegad, 2016, pp. 5157–5164. doi:10.1109/CVPR.2016.445.

[67] J. Shotton, R. Girshick, A. Fitzgibbon, T. Sharp, M. Cook, M. Finocchio, R. Moore, P. Kohli, A. Criminisi, A. Kipman, A. Blake, Efficient human pose estimation from single depth images, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (12) (2013) 2821–2840. doi:10.1109/TPAMI.2012.241.

[68] S. Seyedzadeh, F. Pour Rahimian, P. Rastogi, I. Glesk, Tuning machine learning models for prediction of building energy loads, Sustainable Cities and Society 47. doi:10.1016/j.scs.2019.101484.

[69] S. Seyedzadeh, P. Rastogi, F. Pour Rahimian, S. Oliver, I. Glesk, B. Kumar, Multi-Objective Optimisation for Tuning Building Heating and Cooling Loads Forecasting Models, in: 36th CIB W78 2019 Conference, Newcastle, 2019.
URL https://bit.ly/36IWQIs, Accessed 7th Nov 2019

[70] A. Atapour-Abarghouei, T. P. Breckon, Extended Patch Prioritization for Depth Filling Within Constrained Exemplar-Based RGB-D Image Completion, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 10882 LNCS, Springer, 2018, pp. 306–314. doi:10.1007/978-3-319-93000-8_35.

[71] S. Gupta, R. Girshick, P. Arbelaez, J. Malik, Learning rich features from RGB-D images for object detection and segmentation supplementary material, in: Computer Vision - Eccv 2014, Pt Vii, Vol. 8695, Springer, 2014, pp. 345–360. doi:10.1007/978-3-319-10584-0_23.

[72] N. Silberman, D. Hoiem, P. Kohli, R. Fergus, Indoor segmentation and support inference from RGBD images Lecture Notes in Computer Science, in: ECCV'12: Proceedings of the 12th European conference on Computer Vision - Volume Part V, Vol. Part V, Springer, 2012, pp. 746–760. doi:10.1007/978-3-642-33715-4_54.

[73] J. McCormac, A. Handa, S. Leutenegger, A. J. Davison, SceneNet RGB-D: Can 5M Synthetic Images Beat Generic ImageNet Pre-training on

Indoor Segmentation?, in: Proceedings of the IEEE International Conference on Computer Vision, Vol. 2017-Octob, 2017, pp. 2697–2706. doi:10.1109/ICCV.2017.292.

[74] C. Panagiotakis, H. Papadakis, E. Grinias, N. Komodakis, P. Fragopoulou, G. Tziritas, Interactive image segmentation based on synthetic graph coordinates, Pattern Recognition 46 (11) (2013) 2940–2952. doi:10.1016/j.patcog.2013.04.004.

[75] G. Ros, L. Sellart, J. Materzynska, D. Vazquez, A. M. Lopez, The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2016-Decem, 2016, pp. 3234–3243. doi:10.1109/CVPR.2016.352.

[76] C. Hazirbas, L. Ma, C. Domokos, D. Cremers, FuseNet: Incorporating depth into semantic segmentation via fusion-based CNN architecture, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 10111 LNCS, Springer, 2017, pp. 213–228. doi:10.1007/978-3-319-54181-5_14.

[77] Unity Technologies, Unity.
URL https://unity3d.com/, Accessed 7th Nov 2019

[78] S. Oliver, S. Seyedzadeh, F. Pour Rahimian, B. Kumar, A BIM-centric convergent realities framework for constructing self-propagating hierarchical convolutional neural network ensembles, in: Accepted for presentation at 19th International Conference on Construction Applications of Virtual Reality, Bangkok, 2019.

[79] L. Engstrom, Fast Style Transfer in TensorFlow (2016).
URL https://github.com/lengstrom/fast-style-transfer, Accessed 7th Nov 2019