An Automated Portfolio-based Strategic Asset Management Approach Based on Deep Neural Image Classification

Zigeng Fang[a] Tan Tan[a] Jiayi Yan[a] Qiuchen Lu[a] Michael Pitt[a] Sean Hanna[b]
[a] The Bartlett School of Sustainable Construction, University College London, London, UK
[b] The Bartlett School of Architecture, University College London, London, UK

Image-based classification techniques have historically been applied to identify different building materials in construction sites. However, it was only in recent years that researchers started to explore their applications in built assets' portfolio-based strategic asset management (SAM). Some studies test the feasibility of the image-based technique for the inspection-repair process in a showroom setting with a small image category number. However, real-world healthcare or educational projects would contain more than twenty thousand different assets with over one hundred different asset categories that come from different project portfolios. Currently, there is no available portfolio-based large-scale inspection-repair dataset in operation and maintenance (O&M) phase that was built on images based on the real workflow of building and infrastructure projects. Furthermore, the current image classification application in the building and infrastructure projects' O&M phase has only been applied for object type identification, and not the rest of portfolio-based SAM tasks. Five knowledge contributions are proposed by this study to address the above issues. Firstly, this study proposed a MobileNet-based image classification method for optimising and automating a series of portfolio-based SAM service processes, including: condition surveying process, portfolio-based data validation process, and existing project data's standardisation and integration process. Secondly, a large dataset based on on-site building surveying is collected and built for training and testing, consisting of 11,526 on-site photos (plus 137,530 images crawled from the Google search engine), 274 different built asset categories and twenty-one different projects from ten portfolios. Thirdly, the constructed MobileNet1.0 model achieved satisfactory built assets category classification results based on commercial surveying data, with 72.2% test accuracy with level-3 building cost information service (BCIS) code. Fourthly, it is demonstrated that the MobileNet1.0 model can further boost the on-site image classification performance (a 5.7% Test top 5 accuracy increase) with the enhancement of the online image training dataset. Finally, this study identifies the potential of the MobileNet1.0 model for predicting a series of different building data attributes (e.g., built assets' condition, activity cycle, failure type, residual life, etc.), other than the built assets category. Theese contributions demonstrate the broader application of image-based technologies in portfolio-based SAM and other building and infrastructure projects' O&M phase applications.

Keywords: Convolutional Neural Networks (CNN), Image Classification, Building Condition Survey, Strategic Asset Management (SAM), Operation & Maintenance (O&M)

1 INTRODUCTION

Strategic Asset Management (SAM) is defined as a series of services supplied during the building and infrastructure project's operation and maintenance (O&M) phase that help organizations' and stakeholders' decision-making process related to costs, risks, opportunities of assets throughout the whole lifecycle (Standardization, 2014). Recently, the rising needs for portfolio-based project management have made intelligent asset management requirements even more complex and demanding (Fang et al., 2022). The extended requirement like the integration of different projects' asset information into a centralised strategic "data pool" requires more automated solutions in both the innovation in process and technology. Despite the fact that the number of academic works related to SAM or portfolio-based SAM in the architecture, engineering, construction and facilities management (AEC/FM) sector is limited (Too and Too, 2010), its values and significance cannot be neglected in practices, especially for capital intensive industries, including infrastructure organizations (Gavrikova et al., 2020). The success of decision-making within the SAM is determined by whether the required data can be collected and integrated in an accurate and efficient manner. Moreover, managing the portfolio-based built assets in a strategic manner mitigates potential health and legal risks, and elevates the total asset value of projects from the portfolio level. As one of the pillars for supporting the portfolio-based SAM built asset lifecycle management service, the information required for condition-based maintenance and its major data collection method building condition survey plays a very important role in the portfolio-based SAM services.

Condition-based maintenance has long been regarded as a maintenance strategy bringing efficiency improvement and cost-saving benefits (Prajapati et al., 2012), which is an essential component of the SAM. The built assets condition survey is currently the industry preferred way of collecting data that support the condition-based maintenance service. This is applied on a regular basis to evaluate whether the plant equipment and fabric installation can continually support the functional operation in the whole life cycle. Timely remedial actions can thereby be taken to meet their desired condition level, i.e., modifications, refurbishment, and replacements. Moreover, because building and infrastructure assets' physical condition and operational performance constantly deteriorate with time and use, which results in the depreciation of the built assets' monetary value, knowing the asset condition is important, from the financial perspective, in issuing a strategic asset management plan for stakeholders.

Unfortunately, the conduct of built assets condition surveys is often hindered because of both the lack of technical personnel and financial resources support, which seems to be acknowledged as a worldwide issue in the industry (Maeda et al., 2018). These barriers lead to negative outcomes like low efficiency in the condition survey process and deficiency in data richness and accuracy. In addition, collecting condition survey data in a consistent and standardized manner has also remained a problem that has not been solved during the portfolio-based SAM data collection and management processes. These limitations of current practice suggest the need for an effective technical method to improve the efficiency, accuracy, and standardization of built assets condition surveys, as well as the other portfolio-based SAM processes.

Image classification shows great promise for improving the built asset condition survey and other portfolio-based SAM service processes. Unfortunately, few papers have been published in FM related image classification areas, and most cover onlya pilot stage. For instance, (Zhan et al., 2019) conducted a pilot study to use the image classification algorithm and quick response (QR) code technique to improve the BIM integrated inspection-repair process, but the dataset is limited to sixty images for six different home appliance classes regarding a bedroom and kitchen example space (Zhan et al., 2019). (Marzouk and Zaher, 2020) have developed an automated mechanical, electrical and plumbing (MEP) elements classification and localization system using the simple AlexNet, but the dataset contains only three object categories

(Marzouk and Zaher, 2020). These pilot-level studies' datasets are too small to validate the applicability of image classification for the real built assets' O&M applications, and are often built based on a single project. Research works of applying the image-based algorithm in a larger scaled data set using the actual photos taken by practising surveyors are needed.

These existing studies do not address how image-based classification can amend the enormous practical gap in built assets condition survey, where the efficiency of workflow, object classification accuracy, and standardization of collected data should also be focused and improved accordingly. Multiple problems of human interpretation exist in practice. Firstly, when surveyors enter the asset condition information of different objects, they need to choose from an over-long drop-down list containing over a hundred categories (e.g., "Ceiling", "Floors", "Furniture-fixed", etc.). This rigid process lowers the efficiency of the condition survey greatly. Secondly, it is common in practice that different surveyors from junior to senior level label the same asset with different object categories (e.g., "Exterior-doors", "External-doors", "Internal doors"). If an image-based classification can be adopted, an automated suggestion of category can then be offered to shorten and improve the accuracy of the selection process. The knowledge gap of surveyors with different experiences can be reduced, especially for some junior level surveyors. Moreover, stakeholders may wish to transfer the 'knowledge' and 'wisdom' from particular projects to the portfolio-level strategy (Mohammed, 2021). An image-based classification method which can generalise condition survey data across multiple projects at the portfolio level can do so to benefit each individual project. If a high level of efficiency, accuracy, and standardization of data collection can be achieved in the O&M phase (i.e. the built assets condition survey), by the adoption of image-based classification techniques, the performance of SAM can be optimized.

To do so, an exploration of image-based classification in practical scenarios is urgently needed. Considering the abovementioned problems and gaps, in this study, we explore the applicability and potential of an automated deep learning-based image classification method for portfolio-based strategic asset data collection and management. The contributions of this study are as follows.

1. This study developed an automated convolutional neural network (CNN) based MobileNet image classification model solution for a series of portfolio-based SAM service processes regarding different built assets data attributes. This allows a series of portfolio-based SAM service processes, like the condition surveying process, portfolio-based data validation process, and existing project data's standardisation and integration process to be further optimised and automated.

2. The study collected and prepared a large size building surveying dataset with the real-world setting, containing 11,526 on-site photos and 137,530 images crawled from the Google search engine. 274 different built assets categories from ten portfolios' twenty-one different projects were included.

3. The study demonstrated that the MobileNet1.0 model can provide a satisfying on-site image classification performance (72.2% test accuracy with level-3 BCIS code) based on the commercial projects' surveying image data.

4. The study proved that the MobileNet1.0 model can further boost the on-site test photos' image classification performance (with a 5.7% Test top 5 accuracy increase) given the help of collected strengthening online image training dataset.

5. The study identified the potentials of the developed MobileNet1.0 model automation solution to predict built assets categories' data attributes and the other relevant asset data attribute types.

The rest of the paper is organized as follows. In Section 2, the FM-related image classification studies and datasets were first reviewed with summarised characteristics. Then a brief history of the convolutional deep neural networks for image classification is reviewed with a detailed explanation of why the MobileNet1.0 model is selected for this study and the characteristics of other candidate models. In Section 3, the experiment design is demonstrated. Firstly, the details of how portfolio-based SAM's different services processes are automated and optimised by image classification are discussed. Then the methods and details of how the image dataset for this study is constructed are explained, containing characteristics and background information of on-site and online collected datasets. Section 4 starts with an explanation of detailed experiments and model settings, followed by the validation experiment results and the discussion of the major findings of this study. Finally, this paper is summarised in the concluding section 5.

2 LITERATURE REVIEW

2.1 The trend of image classification in AEC/FM

Image classification is not a new topic in the AEC/FM industry. Many image-related studies have been carried out in recent years, focusing on damage detection (Maeda et al., 2018, Cha et al., 2017, Abeid Neto et al., 2002, Nishikawa et al., 2012, Zalama et al., 2014, Yeum and Dyke, 2015), progress monitoring (Han and Golparvar-Fard, 2015, Hui et al., 2015, Luo et al., 2018), and material recognition (Zhu et al., 2010, Dimitrov and Golparvar-Fard, 2014, Rashidi et al., 2016, Han and Golparvar-Fard, 2015, Bell et al., 2015). However, only a few papers have been published in FM related image classification topics (Zhan et al., 2019, Marzouk and Zaher, 2020) in the last few years. Most of them stayed only at the pilot stage and were tested over a demo-based single project.

Image-based classification techniques have been proposed (Lu and Lee, 2017) as a possible remedy for previously raised issues regarding built assets condition surveys and other portfolio-based SAM services. The number of papers published in FM's O&M phase is limited. (Zhan et al., 2019) applied both the image classification algorithm and quick response (QR) code technique to improve the BIM integrated inspection-repair process information flow with a demo image dataset consisting of sixty images for six different home appliance classes regarding a bedroom and kitchen example space (Zhan et al., 2019). Although the use of QR code and image classification for improving BIM integrated inspection-repair process information flow is innovative, the limited number of built assets covered in the test room cannot effectively evaluate the real applicability of image classification when facing a few hundred different built assets. (Marzouk and Zaher, 2020) attempted to develop an automated mechanical, electrical and plumbing (MEP) elements classification and localization system using the simple AlexNet with a three classes element categories demonstration demo (Marzouk and Zaher, 2020). Again, the limited number of built assets category covered restrain the proving of image classification's application validity in practice. A similar study conducted by (Pezzica et al., 2019) covered many library buildings. Unfortunately, only two were useable for the built assets classification of the four categories selected (including turret, tower, grille exterior, and grille interior). These studies all remained at the pilot stage, where only datasets of images with few categories were used or a single project-level demo scenario was testified. None apply an image-based algorithm in a larger scaled and multi-categorised data set using actual survey photos.

2.2 Image classification datasets for AEC/FM built assets

Overall, there is a lack of large-scale on-site surveying and built assets related datasets with a real-world setting. Privacy issues exist, as built assets elements are full of facility occupants' lives and work, so their related pictures appear in not only the AECFM specific dataset but also the general-purpose dataset. Although many well set-up datasets are available, there is no existing dataset that can be used as the direct training dataset for the building-element-based facility management image classification. Although more and more general-purpose datasets contain fine-grained categories, they still cannot be directly used as the dataset for specialised FM image classification. Therefore, a customized image dataset is needed to enable the experiment required for this study.

In CV fields, general purpose image-based datasets have been constructed to contain the common photos or scenes (Zhu et al., 2019, Zhou et al., 2017, Xiao et al., 2010); however, these were not specially designed for a specific field, such as the AEC industry. Many existing general-purpose datasets suffered from one of the common issues: the lack of FM or building-related categories when utilising them for AECFM specific tasks further. Even for the very large general purposed datasets like "ImageNet" (Deng et al., 2009), which potentially leaded to more AECFM application scenarios and insights, unfortunately, still cannot be used directly as the training dataset for actual FM applications. Although some of the categories were already fine-grained in these giant datasets, they were still too general and not grained in the way that can be used in a real project. The currently available AEC/FM related categories are mainly concentrated on soft FF&E equipment. In addition, some general-purpose datasets have been specially designed and constructed to combine both object detection and semantic scene labelling. These datasets are potentially helpful for other FM related applications like auto-recognition of survey location; however, because this research mainly focuses on the built assets's classification, they do not fit the purpose of this research.

Compared with the general-purpose datasets, AECFM specific datasets are much rarer. Unfortunately, there is no well-established FM related condition survey usable image dataset. For example, the Indoor Scene Recognition dataset (Quattoni and Torralba, 2009) with 67 indoor categories and a total of 15,620 images only contained labels for scenes like bakery, library, gym, etc., which means it cannot be used as the training dataset for building elements' classification. The other two datasets that were found specially designed for AECFM applications are the Construction material Library (CML) dataset (Han and Golparvar-Fard, 2015) and the Materials in Context Dataset (MINC) dataset (Bell et al., 2015). From their names, it is not hard to recognize that they are both image datasets about the building materials. However, the category size of the MINC dataset was still limited (23 categories in total), and the material was not bound to the building element, which made it impossible to be used in the real FM application. The remaining two AECFM related datasets (BIM knowledge repository, Zhan et al., 2019; FM200, Marzouk and Zaher, 2020) only contained 60 and 81 images for the training, which was not sufficient. A new dataset needs to be created to support the FM-based ML applications further. In addition, a summary table (Appendix 1) has also been made to show the comparison between these different datasets, including their major characteristics and statistics related to built assets image classification applications.

2.3 Building condition survey, assessment and management

For any building or infrastructure, degradation is inevitable due to various reasons like day-to-day usage, weather, and erosion resulting from inadequate maintenance (Faqih and Zayed, 2021b). Building assessment is needed and performed by corresponding experts to visually inspect the different building systems (e.g., architectural, electrical, and mechanical) to estimate the current state and extent of deterioration (Bernat and Gil, 2013). The building or infrastructure condition survey is a detailed physical check over the health of a building or infrastructure conducted typically by a qualified building/mechanical surveyor (Chima et al., 2021). Building condition assessment is important as the data collected through the building condition survey forms the basis of the portfolio-based SAM processes and will majorly impact the build assets' lifecycle management decision (e.g., replacement or repair). If unnecessary or improper maintenance activities were carried out due to the insufficient or poor management of built assets information, it would cost up to one-third of all maintenance costs (Mobley, 2002). The accurate identification of building defects through condition assessment or survey could help extend the service life of existing building elements to reduce the emission of greenhouse gas (Paulo et al., 2014).

Unfortunately, many problems exist in the existing condition survey/assessment/management processes. The lack of objectivity and accuracy and time-consuming building condition assessment processes are three major obstacles for built assets managers (Faqih and Zayed, 2021a, Ferraz et al., 2016, Faqih and Zayed, 2021b). There are two major reasons behind this: (1) the current condition assessment process is dependent on the experience and the training of the surveyors and inspectors and (2) a large portion of surveyors' time has been spent on repetitive tasks, like writing notes, that do not demand an expert-level knowledge. The built assets surveying and assessing process needed a support mechanism and technology assistance, like image classification, to assist in achieving more efficient and subjective differentiation among different asset categorical information (i.e., asset condition information (good/fair/poor/critical)) (Mayo and Karanja, 2018). Similarly, the existing spreadsheet-based way of built assets data management is error-prone and tedious. An automated way of error checking was also needed to validate portfolio-based SAM data (Faqih and Zayed, 2021b). Furthermore, for the portfolio-based SAM, there are also needs to standardise the built assets to achieve better project data integration and avoid manual operation.

2.2 Deep learning-based image classification networks

Traditional machine learning algorithms like SIFT began to address the image classification task in the late 1990s. In 2012, Krizhevsky et al. (2012) published the first deep learning model, which set a superior performance benchmark for the task, after which researchers moved on to focus their research on creating deeper and more complex convolutional models. CNN models have outperformed the traditional models in many of the 2-D based CV tasks (e.g., image classification) (Russakovsky et al., 2015). CNN models had also been the basis for other computer vision tasks such as localization, detection, segmentation, and simultaneous localization and Mapping (SLAM) (Karpathy, 2016). The further introduction of transfer learning is proved to boost the classification accuracy of the CNN model algorithm to the next level (Shaha and Pawar, 2018).

2.2.1 Inception Net

Soon after the success of AlexNet (Krizhevsky et al., 2012), the crown of the ImageNet Challenge: ILSVRC (Russakovsky et al., 2015) was taken by its second major successor GoogLeNet (InceptionV1) (Szegedy et al., 2015), which used a new network organization form named "Inception module", which significantly improved the model's performance. Szegedy et al. (2016) further adapted the model to the V2 version to further improve the efficiency of the Inception Model V1. The GoogLeNet's latest version – Inception V4, inspired by the ResNet He et al. (2016a)'s creative structure design, had also included the "Short-cut Connection" further to control the width and height of the grid module and model brunches to improve the network's performance further.

2.2.2 VGG

The VGG model (Simonyan and Zisserman, 2014) also directly led to further improvement of many state-of-art networks (e.g. Inception V2 (Szegedy et al., 2016)). One of the primary aims of Simonyan's research was to evaluate how convolutional network depth will impact its accuracy in the large-scale image recognition setting. Moreover, Simonyan's study further suggested that the Local Response Normalization used in the AlexNet (Krizhevsky et al., 2012) doesn't improve the performance of the convolutional network.

2.2.3 ResNet & SENet

Deeper neural networks are generally more difficult to train. When more layers were stacked, the problem of vanishing/exploding gradients occurs (Bengio et al., 1994, Glorot and Bengio, 2010). Luckily, this difficulty has been addressed by using batch normalization (or normalized initialization) so that these "deep" networks can start converging for stochastic gradient descent (SGD) with backpropagation. Unfortunately, the degradation problem has then been exposed, which was not caused by the overfitting. In this case, adding more layers to the already deep network only causes a higher training error (He et al., 2016a). He addressed this problem by introducing a deep residual learning framework. The same year, He et al. (2016b) further improved the network (ResNet (V2)) by using the "pre-activation" (ReLU and BN (Ioffe and Szegedy, 2015)) of the weight layers, rather than the traditional "post-activation" to make further the information path cleaner. Furthermore, based on the ResNet, SENet Hu et al. (2018) came up with a new mechanism called channel-attention, which can continuously readjust the channel feature responses and provide competitive performance against the ResNet.

2.2.4 DenseNet

Before DenseNet, many different networks had attempted to train end-to-end networks with huge layer sizes. DenseNet innovated by connecting all layers (with the matched feature-map size) directly with each other (Huang et al., 2017). In each dense block, additional inputs from all proceeding layers obtained by each layer pass its feature maps to all subsequent layers. Where traditional convolutional networks have N connections for N layers, DenseNet has

N(N+1)2 direct connections instead. In other words, there is one connection between each layer and its subsequent layer.

2.2.6 SqueezeNet

Unlike some of the previous networks that focused on accuracy improvement (Szegedy et al., 2016, He et al., 2016a). SqueezeNet (Iandola et al., 2016) was developed to achieve comparable accuracy with much fewer parameters. This small-sized model tended to have a faster training speed due to the less communication overhead proportional to the model's parameter number (Iandola et al., 2016). The network used the Fire module consisting of a 'squeeze' convolutional layer (which has 1×1 filters only) and an 'expand' layer consisting of a mix of 3×3 and 1×1 convolution filters to enlarge the activation maps so as to achieve a higher classification accuracy with the reduced input channels.

2.2.5 MobileNet

The success of the AlexNet (Krizhevsky et al., 2012) drove a trend of more complicated and more in-depth networks to push accuracy even further. However, this advancement in accuracy was not always efficient, especially regarding the networks' size and speed. MobileNet (Howard et al., 2017) was developed to provide a timelier and more "portable" network for real-world applications like automatic car driving, robotics and augmented reality. For facility management, it means some original only desktop or stationary workstation operatable tasks can be carried off-line in a portable device like an iPad. The secret weapon used by the MobileNet is a technique called depthwise separable convolution, which factorises a standard shaped convolution into a depthwise convolution.

2.2.6 The proposed model for this study - MobileNet

After comparing the characteristics of different deep neural networks, MobileNet1.0 was chosen as the proposed model for this study. The image classification network needed to be embedded inside the tabular-based surveying apps (e.g., Asseticom and Kykloud) and usually operated off-line in the mechanical room or concealed places with no Wi-Fi signal. Therefore, it is important that the network is compact enough to be deployed given the limited tabular- or mobile- based devices' memory and calculation power; MobileNet achieves this. It is built based on a streamlined architecture that uses depth-wise separable convolutions to enable a lightweight deep neural network structure (Howard et al., 2017). The two simple global hyper-parameters that were introduced provide developers with more flexibility over choosing the right sized model for their applications based on the corresponding constraints. MobileNet's small and low latency model makes it the perfect candidate for mobile and embedded vision applications (Howard et al., 2017). Given the similar level of object detection rate, the MobileNet structure only requires half of the number of parameters needed for the Inception V2 model (Szegedy et al., 2016) and one-fifth of the VGG model (Simonyan and Zisserman, 2014).

3 THE DEVELOPMENT OF THE AUTOMATED PORTFOLIO-BASED SAM APPROACH

3.1 The portfolio-based SAM approach development

3.1.1 *Optimizing surveying processes by image classification*

In the current built assets condition survey process, tabular-based surveying (i.e., Asseticom (Asseticom, 2022) and Kykloud (Gordian, 2022)) through mobile devices is used partially to collect assets information. In addition to the text records of asset condition, image is an important part of the condition survey as it serves as the evidence of showing the building asset condition. While images bring positive feedback on providing valuable asset condition data, a potential efficiency improvement gap has been observed in a practical scenario of the condition survey process. Specifically, when data entries of built assets are input respectively, surveyors are responsible for selecting the corresponding object categories from a very long drop-down list consisting of over a hundred categories through the mobile device (shown in Figure 1). As a result, the surveyor needs to spend a lot of time struggling with the manual category labelling work. And the situation is even worse for junior surveyors due to the unfamiliarity with building asset categories. This manually conducted rigid process introduces side effects that either the time spent on the multiple built assets condition survey increases (which leads to over-budget) or the details of survey data collection are omitted.

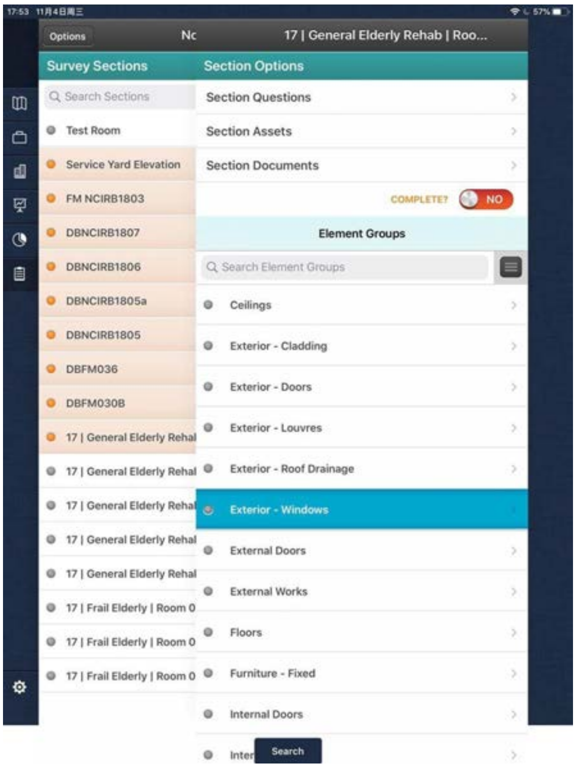


Figure 1 The screen shot showing the object categories' long drop-down list in a tabular-based surveying software (Gordian, 2022)

These side effects might be eased by image-based classification, by providing a shortlist of suggested object categories automatically after the survey image is taken. This change of condition survey workflow improves the efficiency of the entire surveying process while promoting data consistency by reducing the deviation range of manual selection, with positive impact on survey time and human resources costs in addition to input data consistency.

3.1.2 *Optimizing surveying data validation processes by image classification*

Image classification may also improve the built assets condition survey through data validation. The current management and maintenance process of collected built assets data is still heavily dependent on manual work, which introduces several errors and time-consuming problems (Fang et al., 2022). First, since the surveyors' expertise and experience levels vary greatly, different surveyors seem to have different category labelling preferences. This causes data inconsistence and inaccuracy across the existing projects during the data management process, especially from a portfolio-level management perspective. Secondly, many necessary assets condition attributes were left blank during the condition survey process (e.g., "Condition", "Activity Cycle", "Unit", "Health and Safety Issue", "Maintenance Issue", "Manufacturer", "Lifetime Source", etc.) by on-site surveyors, which necessitated additional subsequent labour in recategorizing, cleaning and querying data. These problems persist in data down-chain (e.g., COBie files (exported from BIM), CAFM system data, and room datasheet) : once the deficient data becomes integrated into the master portfolio-based SAM model, more serious errors can be made during the decision-making process.

The authors suggest that image-based classification can provide the portfolio-level data with an automatic way of checking the overall quality by offering and filling asset condition information automatically without dependence on manual inputs by surveyors via mobile devices or excel spreadsheet. Furthermore, it validates and enriches the various types of project existing data and survey data collected during the condition survey process.

3.1.3 *Optimising existing (surveying) project-based data integration processes with image classification*

SAM capability has traditionally been limited by single-project-based management. To achieve the goals of the total-best performance of built assets, the portfolio-based innovative SAM has been recognised and adopted. Leveraging the knowledge and information gathered in the master portfolio-based SAM model, the procurement decisions can be compared across the different projects' suppliers for an optimal option and enjoy a larger discount for bulk procurement for the entire portfolio. However, to provide various types of portfolio-based SAM enabled services, data must be standardised and cleaned before transitioning from the project-level to the portfolio-level data model.

Given automated image classification, existing project data from either CAFM system, previous condition survey, or room datasheet can be categorised in the standardised format used for the portfolio-based SAM data model before integrating into the master portfolio-based SAM data model. This avoids the "language barrier" among each project in the same portfolio and ensures the following portfolio-based SAM services (e.g., strategic procurement) can be conducted effortlessly.

3.1.4 *A summary of image classification's integration with SAM portfolio-based detailed services documentation flow structure*

The main idea for using image classification is to mitigate the "lack of information" issue currently faced by SAM and facility managers. In Figure 2, the specific processes are summarised for how image classification is used to help SAM project and portfolio managers to speed up and improve the accuracy of data collection surfing the surveying process, semi-automatically validate the existing portfolio-based data entries' data quality, and provide automated built asset categorisation and standardisation before the different types of existing project data are integrated to a master portfolio.

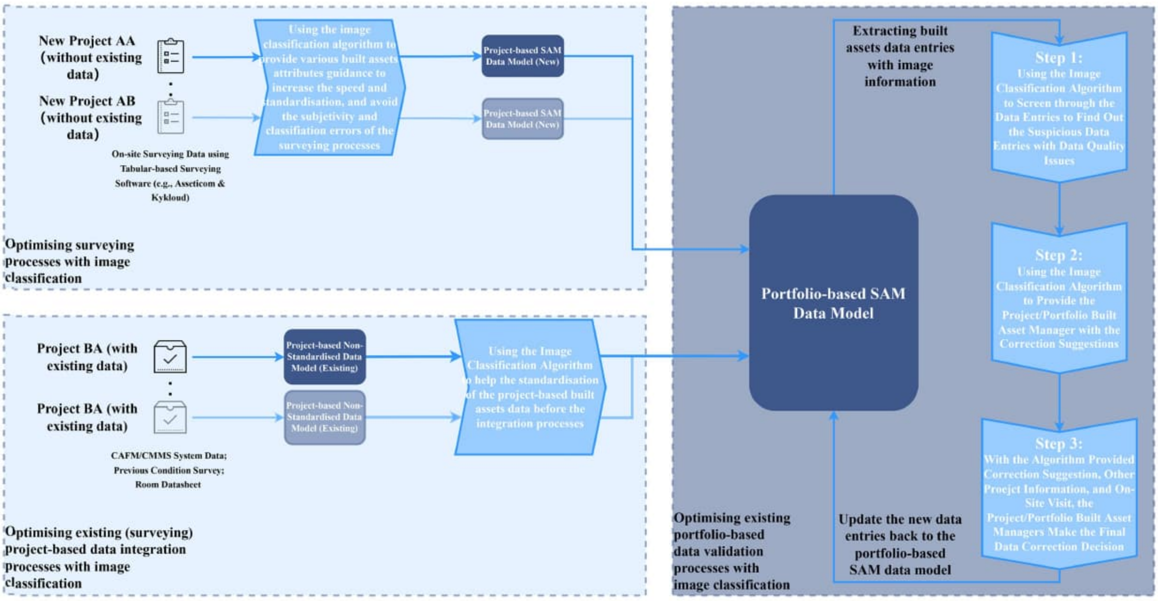


Figure 2 The integration of portfolio-based SAM data model and image classification technology

Image-based classification is the key technique that fastens survey data collection from a new building or infrastructure into the Master SAM Data Model since it guarantees data accuracy and standardization. It also helps with the data handover from the design and construction phases into the operation and maintenance phase in real-life cases, because there is usually a tremendous amount of image data for built assets in the handover datasets that are hard to document and manage manually, let alone integrate the data into the Master SAM Data Model. The image-based method offers a possibility to classify and reuse those data automatically as early-stage asset condition proof data. In general, without the assistance of the updated machine learning image method, the data validation and management at the portfolio level could be mission-impossible during the life-long SAM process, which affects the strategic decision-making process irreversibly and irreparably. Furthermore, portfolio-based management enables the possibility of involving and concatenating a series of projects that share the same strategic goals and competing for the same resources, which helps portfolio and project managers achieve their goals to lower the cost and risks for their clients.

3.2 Proposed dataset establishment

3.2.1 *Data collection*

3.2.1.1 <u>Building element coding structure – customized New Rules of Measurement (NRM) (Data category)</u>

New Rules of Measurement (NRM) is a series of standardized measurement rules developed by RICS that stakeholders can easily understand (e.g., employers, project/design team, surveyors/ cost managers) inside a building and infrastructure construction projects. There are three sections: order of cost estimating and cost planning for capital building works, detailed measurement for building works, and order of cost estimating and cost planning for building maintenance works (RICS, 2012). The aims of NRM are aiding the communication among the stakeholders by eliminating the language barriers and providing reliable advice of cost for all the stakeholders (RICS, 2012).

In this study, NRM plays an important role as the standard of attribute category labelling during the data collection process. The reasons for adopting NRM are (1) NRM series are intended for relevant activities and users (e.g., surveyors) in the O&M phase, which indicates a more targeting and accurate category classification (RICS, 2012); (2) it works as a mediator to break the data obstacle when SAM professionals try to aim for portfolio-based strategic asset management; and 3) although it is developed in the UK, it is applicable for worldwide projects. This allows for the spread of this proposed method worldwide (RICS, 2021).

3.2.1.2 Image data quality analysis (on-site)

3.2.1.2.1 Analysis of main built assets type category images profile (on-site)

In this research, all the on-site image data collected belongs to Vercity UK Ltd's Strategic Asset management Team's project portfolios based in the United Kingdom, most located in England. In total, there are ten project portfolios and twenty-one sub-projects. Overall, the number of projects in the Greater London region and the South west region (both with seven projects) outweigh the rest of the regions. The rest of three regions, the Northwest, the West Midlands, and the South East, each have two projects. In terms of the Gross Internal Floor Area (GIFA), although this cannot be found in some of the portfolio document files, the sizes vary from 5,901 to 83,337 m^2. Within the targeted project portfolios, six are healthcare project portfolios, and three are educational project portfolios; only one project portfolio belongs to the social housing project type. The total lifecycle costs expected for targeted project portfolios' Private Finance Initiative (PFI)/Public-Private Partnership (PPP) concession periods vary from £ 2,300,000 to £ 100,000,000. Not surprisingly, educational and housing projects have relative lower lifecycle expenditures, while the largest five lifecycle expenditure estimations are all come from Healthcare projects. All detailed project portfolio information is summarized in Appendix 2. During the cleaning process of the on-site image dataset, it is suggested that not all the on-site collected images are useable. As a result, a small portion of photos that miss the proper image labels and do not fulfil the basic technical requirement (e.g., a photo that was wrongly taken with the main object missed) are excluded from the final experiment, leaving 11,526 on-site images for the final experiment.

By default, RICS's level-2 and level-3 NRM codes are available for some existing and previous contracted projects. The case study SAM team has developed a more detailed customized level-4 coding structure to meet their projects' demand. By comparing the image distribution profiles in level-2 and level-3 NRM code, it is clear that the largest image category comes to 'Wall

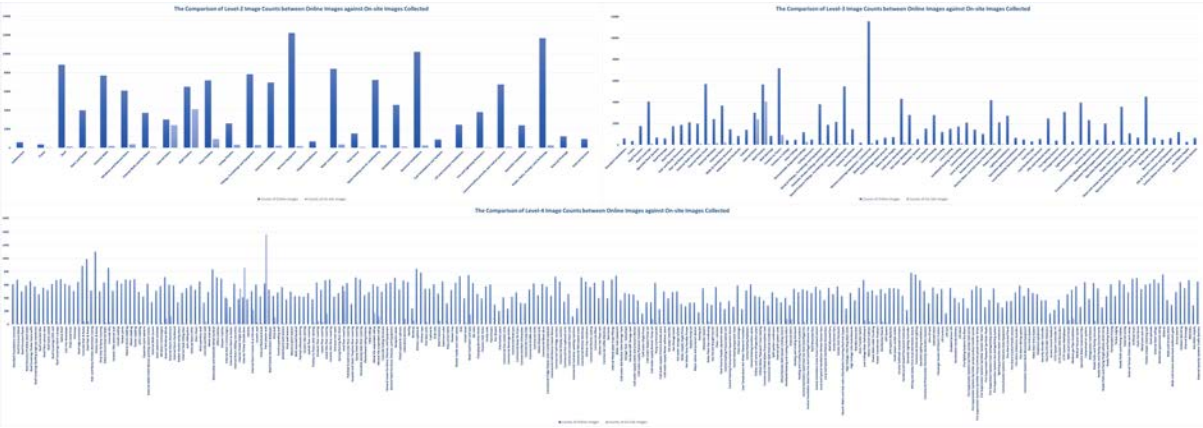


Figure 2 the Number of Different Images with Level-2 BCIS Code Available (Upper Left); the Number of Different Images with Level-3 BCIS Code Available (Upper Right); the Number of Different Images with Level-4 BCIS Code Available (Bottom)

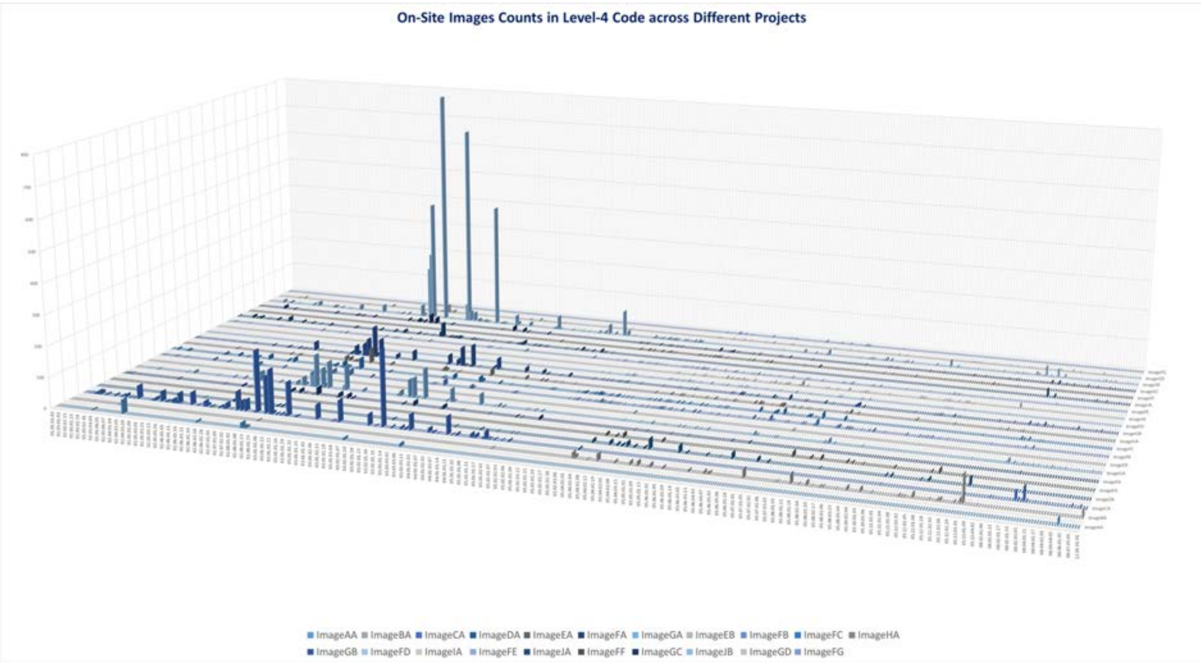


Figure 3 The On-site Images Counts in Level-4 Code Across Different Projects

Finishes – 03.02' (Figure 2). It is shown in Figure 3 the on-site image dataset's distribution is uneven. In other words, more photos are taken within some specific categories. The unevenly distributed category list is due to the current way of surveying, wherein a typical condition survey, photos are only required when image evidence is needed to prove the occurrence of a defect and collected as supporting evidence for the approval of the maintenance budget request. As a result, it is suggested that if the training dataset is directly built based on the on-site images available, the model might not be robust enough and potentially have a high risk of overfitting the large-sized categories. Although this traditional way of surveying limited the number of photos taken during the surveying processes, it provides the collected photos with more chances to carry richer information (e.g., asset maintenance condition information) for the built assets than just the asset category information.

3.2.1.2.2 Analysis of other built assets type category images profile (on-site)

Unlike the previous O&M related AEC/FM related image dataset, the containing of different building asset data attributes is a unique characteristic of this dataset collected during the real-world building condition/defect survey workflow. Unfortunately, some of the previously obtained images don't have the corresponding data attributes that follow the same attribute structure required by tabular based surveying software (i.e., Kykloud). Therefore, these projects' data is excluded from corresponding analysis and experiment. Finally, we have 11,163 images with data attributes that fulfil the required format. In addition, during the data cleaning process, invalid or misleading data entries like '0' and 'N/A' in some data attributes are amended with a default value – 'null'. The detailed image distribution for some of the example data attributes is shown in Figure 4.

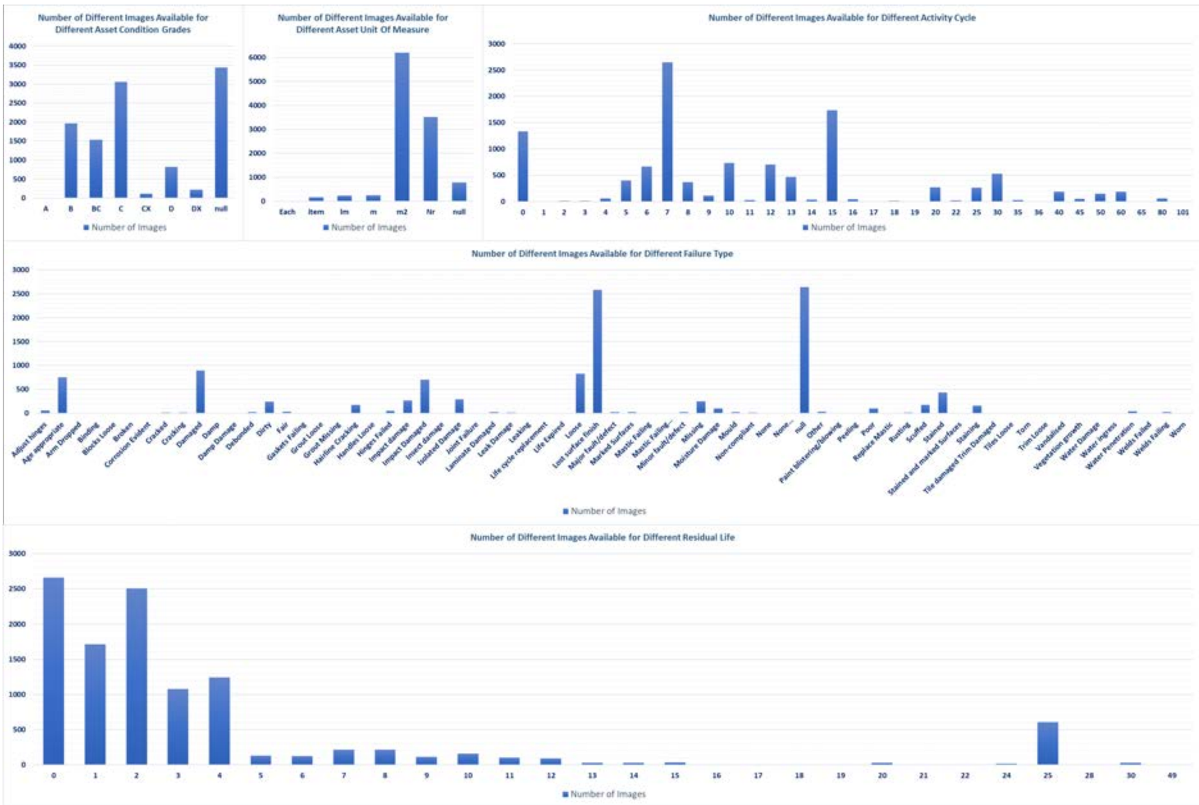


Figure 4 The Number of Different Images with Different Condition Grade (Upper Left); Unit of Measure (Upper Middle); Activity Cycle (Upper Right); Failure Type (Middle); Residual Life (Bottom)

3.2.1.3 Results and statistics of new FM dataset (online)

The previous section demonstrates the statistics of various attached labels of the on-site image dataset's category distribution. However, the size of this research's on-site dataset is already larger than some of the previously mentioned small size datasets (e.g., CML dataset, Han and Golparvar-Fard, 2015) and close to some of the medium-size datasets (e.g., Indoor Scene Recognition, Quattoni and Torralba, 2009) that were designed for AECFM. Unfortunately, due to the on-site dataset's unevenly distributed building asset categories (the SAM Unit's customized BCIS NRM level-4 code, Figure 3), the training based on this dataset alone could potentially suffer from overfitting. Therefore, online images were used to enhance the on-site training dataset with a more evenly distributed training image category distribution.

Construction of the online-based image dataset involved several processes. Firstly, some online downloaded image categories that were too small in quantity or dominated by ambiguous images were eliminated. On average, 43% of the downloaded images are eliminated as they don’t include the target built assets element. Secondly, some keywords for image queries which were not able to find the corresponding images for the building condition survey (especially for the categories that investigate the inside component of the HVAC system: e.g., “Heating & Cooling Coils”) were eliminated. Finally, after synchronizing both on-site and online image datasets’ asset categories, only the categories with identical or directly related asset descriptions were kept. As a result, a total of 274 asset categories (with 137,530 images in total) were chosen as the final categories (Figure 2). Compared with the relatively small image counts per category for the on-site images (29 images per category), the number of online images per category reaches 501 on average. The smallest category in the online image training dataset has more than 100 different images.

3.2.2 *Ethics and privacy*

To protect the privacy of the project portfolio’s involved stakeholders and participants, and to reach the required ethical standard for this research, all the project portfolios and portfolios’ individual project identity information was purposely anonymized and blurred (e.g., ‘ImageA’ for the first project portfolio & ‘ImageAA’ for the first project of the first portfolio, Appendix 2). Furthermore, photos were deliberately taken and rechecked for the on-site photos collection process to avoid capturing humans.

4 RESULTS

4.1 Experiments and models settings

A series of experiments have been conducted to validate the MobileNet’s applicability over built assets condition survey in a real-world scenario. The first experiment tests the performance of different CNN-based image classification networks over the on-site data against MobileNet1.0. The main objective of the first experiment is to compare the performance of different deep neural-based CNN networks upon the large-scale FM-oriented online collected image dataset, to confirm the suitability of MobileNet1.0 in further experiments. The second experiment tests the effect of different image augmentation methods over the on-site image dataset with the MobileNet1.0 model. The objective of the second experiment is to find out the level of improvements that can bring by different image augmentation methods upon the MobileNet1.0 model. The third experiment moves the target task of image classification one step further to predict different built assets data attributes. This experiment explores the potential and applicability of image classification over suggesting other built assets elements’ data attributes. The last experiment targets the measure of improvement over the online image enhanced training dataset over on-site image classification, which aims to identify the level of classification improvement that can be brought by the online enhancement image dataset and validate the MobileNet1.0’s performance against the rest of the two selected models over the real-world on-site testing image dataset.

In this study, the small images diameter of around 225 × 225 (pixels) is chosen as the image size for the data mined images to improve the on-site classification performance. The entire dataset has been shuffled with random seed and segregated into training, validation, and testing datasets (with the ratio of 8:1:1). Due to the limitation of the working platform used in this experiment (in this experiment, an Alienware laptop with NVIDIA GeForce GTX 1080 and 32.0 GB Memory), some models, unfortunately, need to compromise over its parameters (e.g., batch size), as otherwise the program cannot be executed usually due to the system out of memory.

4.2 Image classification training and validation

4.2.1 Validation image classification training results against MobileNet1.0 - upon large size online collected image dataset labelled with different CNN networks

In this experiment, the results of different image classification networks are compared against MobileNet1.0 using the online dataset constructed. Different versions of network structures existed for some networks like DenseNet, which can be distinguished by network parameters size and hyper-parameter settings, were tested. As the major purposeof this experiment is to compare the different models' performances against MobileNet1.0 under the similar condition, same parameter setting is used for most of the models if possible.

In total, eight different model series' fourteen different models are compared. Evaluated for both in both training time and accuracy, MobileNet1.0 offers the most all-round performance with a very competitive test accuracy rate (46.2%), test top 5 accuracy rate (76.4%), and much less training time (7 hours) compared with 20 hours for ResNet152_V1 model. Therefore, the result validates the hypothesis that the proposed MobileNet1.0 model is the most suitable model for this image classification task. The result also shows out of three selected compact network structure candidates (e.g., 1) MobileNet0.25; 2) ResNet18_V1; 3) SqueezeNet1.1), all of them are not computational costly to deploy (with less than 4 hours training time). While the ResNet18_V1 is suggested to be the best compact image classification model structure as it achieves the highest accuracy performance (42.8% test accuracy and 72.8% test top 5 accuracy (Table 1)) among the candidate models with slightly longer training time. Moreover, the result also indicates the most accurate model out of fourteen different model structures is the ResNet152_V1 model, which has the highest accuracy ratings (46.7% test accuracy and 77.6% test top 5 accuracy).

Table 1 Experiment One Table: The Image Classification Result with Online Collected Dataset against Different Network Structures

Model	Test-acc:	Test-top-2-acc:	Test-top-3-acc:	Test-top-5-acc:	Total Time Spent:	Average Time Spent:
AlexNet	36.6%	49.8%	56.8%	65.6%	00:01:44:58	00:00:03:30
DenseNet121	42.8%	56.5%	63.8%	72.7%	00:13:55:39	00:00:27:51
DenseNet161	43.2%	57.3%	64.7%	73.4%	01:06:23:40	00:01:00:47
DenseNet201	34.6%	47.8%	54.8%	63.5%	00:23:45:53	00:00:47:32
InceptionV3	46.0%	60.1%	67.8%	76.1%	00:12:55:02	00:00:25:50
MobileNet0.25	38.6%	51.2%	58.7%	68.0%	00:02:39:01	00:00:05:18
MobileNet1.0	46.2%	61.0%	68.3%	76.4%	00:07:22:40	00:00:14:45
ResNet18_v1	42.8%	57.1%	64.5%	72.8%	00:03:29:22	00:00:06:59
ResNet101_v1	45.5%	60.1%	67.6%	76.0%	00:16:17:35	00:00:32:35
ResNet152_v1	46.7%	61.6%	69.5%	77.6%	00:20:37:47	00:00:41:16
SENet_154	38.8%	52.4%	60.0%	69.1%	04:04:00:02	00:03:20:00
SqueezeNet1.1	31.5%	43.7%	50.6%	60.0%	00:02:03:40	00:00:04:07
VGG11	42.8%	57.3%	64.4%	73.1%	00:06:08:23	00:00:12:17
VGG19	40.4%	54.3%	61.4%	70.0%	00:14:22:45	00:00:28:45

4.2.2 Image classification training results over MobileNet1.0 - upon large size online collected dataset with different image augmentation

In this experiment, different image augmentation methods' effectivenesses are compared against the chosen most balanced model - MobileNet1.0. The result shows that the level of classification performance improvement provided by using different image augmentation techniques is relatively limited. For example, the classification result using random colour (45.3%) and random lighting condition (45.4%) for the training image did not bring significant benefits to the original benchmark result without any image data transformation (45.5%) (Table 2). In comparison, the RandomFlip technique seems to be most effective in improving the test classification accuracy, although the relative improvement is still slightly limited (about 1.3%).

Table 2 Experiment Two Table: The Image Classification Result with Online Collected Dataset against Different Image Augmentation techniques

Model	Test-acc:	Test-top-2-acc:	Test-top-3-acc:	Test-top-5-acc:	Total Time Spent:	Average Time Spent:
Without Data Transformation	45.5%	60.8%	68.3%	76.2%	00:07:31:13	00:00:15:02
RandomCrop	46.2%	61.0%	68.3%	76.4%	00:07:22:40	00:00:14:45
RandomColour	45.3%	60.9%	68.3%	76.3%	00:07:26:47	00:00:14:54
RandomFlip	46.8%	62.3%	69.7%	77.4%	00:07:21:30	00:00:14:43
RandomLightingCondition	45.4%	61.0%	68.6%	76.7%	00:08:35:32	00:00:17:11

4.2.3 Image classification training results over MobileNet1.0 - upon real defect survey dataset labelled with other infrastructure building asset data attributes

The result of the third experiment proves that the image classification cannot only be used for the classification of the built assets category but is also capable of making various types of predictions. Given the different data attributes provided by the on-site collected dataset, it is suggested that the image embedded information can offer a satisfying prediction over almost all of the selected built assets data attributes. This result is supported by referring to the improvements of different data attributes' test classification accuracies against their No Information Rates (NIR), which are the percentage of the largest category, varies from 11.3% for "Percentage of Replacement" to 48.8% for "Activity Cycle" (Table 3). This indicates that the image classification algorithm (MobileNet1.0) can provide helpful built assets data attribute suggestions with the object feature identified within the on-site test photos.

Table 3 Experiment Three Table: The Image Classification Result with On-site Collected Dataset against Different Image Data Attributes

On-site Image Data Attributes	NIR	Test-acc:	Test-top-2-acc:	Test-top-3-acc:	Test-top-5-acc:	Total Time Spent:	Average Time Spent:
Activity Cycle	23.7%	72.5%	83.5%	87.5%	92.5%	00:00:49:21	00:00:01:39
Condition	39.7%	56.3%	77.1%	91.7%	99.5%	00:00:37:23	00:00:01:15
Failure Type Reference	30.3%	76.5%	86.6%	90.6%	95.3%	00:00:41:11	00:00:01:22
Percentage of Parent	66.4%	77.7%	88.3%	92.0%	96.0%	00:00:53:07	00:00:01:46
Residual Life	23.8%	54.4%	68.3%	76.9%	87.8%	00:00:52:40	00:00:01:45

Unit of Measure	59.8%	91.2%	97.8%	99.8%	99.9%	00:00:53:29	00:00:01:47

4.2.4 Validation image classification training results agaisnt MobileNet1.0 – upon combined image dataset or on-site image dataset along

The last experiment results show a noticeable classification performance improvement given the utilisation of both the online image and the on-site image as the combined dataset for training. The classification performances of previously selected representative models (including ResNet18 v1 model, ResNet152 v1 model, and MobileNet1.0 models) have once again reaffirmed the result of the model comparison experiment using the online data (in the first experiment), with the same classification ranking order for different training settings. Although the classification accuracy of the MobileNet1.0 is again not the highest against the rest of the two models, its 60% classification accuracy is satisfied with one third (11 hours) of the total training time taken by the ResNet152 model. It is also indicated that the improvement in test classification accuracy is more significant when more suggestion options are considered (Table 4).

Table 4 Experiment Four Table: The Image Classification Result with Both Online and On-site Collected Dataset against Different Representative Network Structures

Parameter	Test-acc:	Test-top-2-acc:	Test-top-3-acc:	Test-top-5-acc:	Total Time Spent:	Average Time Spent:
ResNet18_v1 On-site & Online	58.5%	75.4%	83.0%	87.8%	00:04:34:00	00:00:09:08
ResNet18_v1 On-site	57.2%	71.6%	77.9%	82.9%	00:00:57:19	00:00:01:55
ResNet152_v1 On-site & Online	62.3%	78.7%	85.9%	90.2%	01:05:27:34	00:00:58:55
ResNet152_v1 On-site	61.4%	74.7%	80.2%	84.7%	00:02:02:12	00:00:04:04
MobileNet1.0 On-site & Online	60.0%	76.8%	84.3%	89.0%	00:11:25:32	00:00:22:51
MobileNet1.0 On-site	57.7%	72.3%	77.9%	83.3%	00:01:10:56	00:00:02:22

Furthermore, the result (Table 5) indicates that the final prediction using combined on-site and online datasets gives a superior classification performance over the test dataset's NIR. An average of 40% prediction improvement is expected for the different levels of the built assets category. This finding has also been supported by the confusion matrix of the test dataset classification results (Figure 5 & 6), as a large number of the predictions fall into the right cells (the diagonal cells of the confusion matrix that are filled with blue colour).

Table 5 The Image Classification Result with Both Online and On-site Collected Dataset over MobileNet1.0 against NIR

Rate (Comparison with No Information Rate (NIR))	Level-1 Category Test-acc:	Level-2 Category Test-acc:	Level-3 Category Test-acc:	Level-4 Category Test-acc:
MobileNet1.0 On-site & Online	85.6%	74.4%	72.2%	57.5%
NIR	42.9%	32.9%	32.9%	10.1%

Original Label		Target Label					Accuracy
		02	03	04	05	08	
02	Superstructure	305	43	6	13	4	82.2%
03	Internal Finishes	26	512	8	4	2	92.8%
04	Fittings, Furnishings and Equipment	4	8	25	0	0	67.6%
05	Services	24	24	3	229	2	81.2%
08	External Works	8	3	1	2	30	68.2%

Figure 5 The Confusion Matrix of Built Assets (BCIS (NRM) Code) Level-1 Code Test Classification Result

Original Label		Target Label																												Accuracy
		02.01	02.03	02.04	02.05	02.06	02.07	02.08	03.01	03.02	03.03	04.01	05.01	05.02	05.03	05.04	05.05	05.06	05.07	05.08	05.09	05.10	05.11	05.12	05.13	08.02	08.04	08.06	08.07	
02.01	Frame	3	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60.0%
02.03	Roof	0	9	1	2	2	0	0	1	0	1	0	0	0	0	1	1	2	0	0	0	0	0	1	1	1	1	0	0	37.5%
02.04	Stairs and Ramps	0	0	7	0	0	0	1	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53.8%
02.05	External Walls	0	1	0	15	2	0	2	5	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	53.6%
02.06	Windows and External Doors	0	1	0	0	26	1	7	2	2	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	61.9%
02.07	Internal Walls and Partitions	0	0	0	0	0	6	1	0	3	0	2	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	40.0%
02.08	Internal Doors	0	1	0	0	5	0	212	9	8	1	4	0	0	0	2	0	1	0	0	0	1	0	0	0	0	0	0	0	86.9%
03.01	Wall Finishes	0	0	0	3	2	1	11	380	11	3	4	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	91.1%
03.02	Floor Finishes	0	0	1	1	0	0	6	9	76	0	3	0	0	0	0	1	1	0	0	0	0	0	0	0	2	0	0	0	76.0%
03.03	Ceiling Finishes	0	0	0	0	1	0	0	7	0	26	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	74.3%
04.01	Fittings, Furnishings and Equipment	0	0	0	0	1	0	3	5	3	0	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	67.6%
05.01	Sanitary Installations	0	0	0	1	0	0	1	5	3	1	1	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	63.6%
05.02	Services Equipment	0	0	0	0	0	0	0	1	0	0	1	1	19	0	2	2	3	2	2	0	1	0	1	0	0	0	0	0	54.3%
05.03	Disposal Installations	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0%
05.04	Water Installtions	0	0	2	0	0	0	4	1	1	1	0	2	1	0	22	3	3	1	2	0	0	1	1	1	0	0	0	0	47.8%
05.05	Heat Source	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	5	3	0	0	0	0	0	0	0	0	0	0	0	45.5%
05.06	Space heating and air conditioning	0	0	0	1	1	0	0	1	2	1	0	0	0	0	5	0	24	1	4	0	0	0	2	0	0	0	0	0	57.1%
05.07	Ventilation Systems	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	2	9	1	0	0	1	1	0	0	0	0	0	50.0%
05.08	Electrical Installations	0	0	0	0	2	1	1	4	0	0	1	0	0	0	1	0	5	1	18	0	0	0	3	0	0	1	0	0	47.4%
05.09	Fuel Installations and Systems	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0.0%
05.10	Lift and Conveyor Installations	0	0	0	0	0	0	2	1	0	0	0	0	0	0	1	0	0	0	0	0	5	0	0	0	0	0	0	0	55.6%
05.11	Fire and Lightning Protection	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	1	0	0	5	0	0	0	0	1	0	45.5%
05.12	Communication,security and control systems	0	0	0	0	1	1	3	0	0	0	0	0	2	0	1	0	1	0	1	0	0	0	9	1	0	0	0	0	45.0%
05.13	Specialist Installations	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	1	1	1	0	0	0	0	10	0	0	0	0	62.5%
08.02	Roads, Paths, Paving's and Surfacings	0	4	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	71.4%
08.04	Fencing, Railings and Walls	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	4	4	0	0	33.3%
08.06	External Drainage	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.0%
08.07	External Services	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	50.0%

Figure 6 The Confusion Matrix of Built Assets (BCIS (NRM) Code) Level-2 Code Test Classification Result

4.3 Evaulation of fourth experiment result with visualization of randomly selected items

To further evaluate the classification results of the main experiment (the fourth experiment) of this study, the image classification results of five randomly selected test images are visualized in Appendix 3 using the most balanced MobileNet1.0 and the rest of two superior classification models (ResNet152 v1 and ResNet18 v1). The highest five-category recommendations are listed for each of these three models with their category names, customized BCIS codes, and relative probabilities. Firstly, using the combined training data offers a more generalized and logical sound prediction than only on-site photos. For instance, in the classification result of ResNet152 v1 model over the fifth image, more evenly distributed confidence probabilities among "Commercial Gas cooker units", "Commercial Combi Oven units", and "kitchen Hot Plate unit" make more sense than a 98% confident "Commercial Deep fat fryers". The visualization of the classification result has again reaffirmed that the prediction of all these three models is logical sound, even for the wrongly predicted images from the categories like "CCTV Camera" and "Commercial 4-way hobs". Taking the prediction of the fifth image (Appendix 3) (from "Commercial 4-way hobs") as an example, the top-3 predictions of MobileNet1.0 for the fifth image are "Commercial Deep fat fryers", "Commercial Hot cupboard units", and "Commercial Gas cooker units", which is logically sound. Although these options did not include “Commercial 4-way hobs”, all of these four level-4 categories belong to the same level-3 group: 05.02.02. In many cases, the appearance of highly similar category groups that overlap each other is unavoidable. In this case, “Commercial Gas cooker units” could be viewed as the replaceable group for the “Commercial 4-way hobs”, because project AA’s “Commercial 4-way hobs” might be registered as “Commercial Gas Cooker Units” in project BA in the real world. The probabilities given for each classification suggestion can benefit junior surveyors by helping them choose the best asset category out of competitive candidates.

These specific classifications (Appendix 3) confirm the results of section 4.2 in that, compared with the ResNet152 v1 model’s more probability concentration on the top suggestion, the MobileNet1.0 and ResNet18 v1 models seem to have the more evenly distributed confidence level.

5 CONCLUSIONS

This study developed a new large-scale built assets surveying dataset. With the help of search engines (i.e., Google) and real-world on-site asset surveying projects, 11,526 on-site photos and 137,530 web crawled images were collected and used as the training and testing datasets. This is currently one of the largest AECFM specific image datasets that has been built for the building and infrastructure project’s operation and maintenance phase. With its unique characteristics of providing various asset information attributes, we think this dataset opens a new research direction for the image-based technique used for the FM purpose. We also tested the performance of a proposed model, using MobileNet1.0, by comparing it with other state-of-art image classification models over the built dataset to determine the best image classification solution for the corresponding tasks, and the current reasonable expectation of error. The result shows the MobileNet1.0 image classification model can provide satisfying (72.2% and 57.5% accuracy rating, compared with 32.9% and 10.1% NIR for customized level-3 and level-4 BCIS NRM Code) building object category classification results in the real building defect and condition survey scenarios. This study provides the future directions of how machine learning techniques like image classification can benefit the portfolio-based SAM services through optimising and automating the condition surveying process, portfolio-based data validation process, and existing project data’s standardisation and integration process. This study demonstrates that machine learning-based techniques can be applied to both project and portfolio-based SAM and other O&M tasks under the building or infrastructure project’s operation and maintenance phase.

REFERENCES

ABEID NETO, J., ARDITI, D. & EVENS, M. W. 2002. Using colors to detect structural components in digital pictures. *Computer‐Aided Civil and Infrastructure Engineering,* 17**,** 61-67.

ASSETICOM. 2022. *Better Asset Condition Surveying* [Online]. Asseticom. Available: https://asseticom.co.uk/ [Accessed 23/01/2022 2022].

BASHIRI, F. S., LAROSE, E., PEISSIG, P. & TAFTI, A. P. 2018. MCIndoor20000: A fully-labeled image dataset to advance indoor objects detection. *Data in brief,* 17**,** 71-75.

BELL, S., UPCHURCH, P., SNAVELY, N. & BALA, K. Material recognition in the wild with the materials in context database. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 3479-3487.

BENGIO, Y., SIMARD, P. & FRASCONI, P. 1994. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks,* 5**,** 157-166.

BERNAT, E. & GIL, L. 2013. Aided diagnosis of structural pathologies with an expert system. *Advances in Structural Engineering,* 16**,** 379-393.

CHA, Y. J., CHOI, W. & BÜYÜKÖZTÜRK, O. 2017. Deep learning‐based crack damage detection using convolutional neural networks. *Computer‐Aided Civil and Infrastructure Engineering,* 32**,** 361-378.

CHALADZE, G. & KALATOZISHVILI, L. 2017. Linnaeus 5 dataset for machine learning. Technical report.

CHIMA, O. A., IFEANYICHUKWU, N. E., NNAMDI, A. T., CHUKWUEMEKA, I. G. & PRECIOUS-FAVOUR, N. O. 2021. Evaluation of the Practice of Building Survey for Maintenance of Public Building in South-East Nigeria. *Journal of Engineering Research and Reports***,** 108-122.

DARLOW, L. N., CROWLEY, E. J., ANTONIOU, A. & STORKEY, A. J. 2018. Cinic-10 is not imagenet or cifar-10. *arXiv preprint arXiv:1810.03505*.

DENG, J., DONG, W., SOCHER, R., LI, L.-J., LI, K. & FEI-FEI, L. Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition, 2009. Ieee, 248-255.

DIMITROV, A. & GOLPARVAR-FARD, M. 2014. Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections. *Advanced Engineering Informatics,* 28**,** 37-49.

EVERINGHAM, M., VAN GOOL, L., WILLIAMS, C. K., WINN, J. & ZISSERMAN, A. 2010. The pascal visual object classes (voc) challenge. *International journal of computer vision,* 88**,** 303-338.

FANG, Z., LIU, Y., LU, Q., PITT, M., HANNA, S. & TIAN, Z. 2022. BIM-integrated portfolio-based strategic asset data quality management. *Automation in Construction,* 134**,** 104070.

FAQIH, F. & ZAYED, T. 2021a. A comparative review of building component rating systems. *Journal of Building Engineering,* 33**,** 101588.

FAQIH, F. & ZAYED, T. 2021b. Defect-based building condition assessment. *Building and Environment,* 191**,** 107575.

FEI-FEI, L., FERGUS, R. & PERONA, P. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. 2004 conference on computer vision and pattern recognition workshop, 2004. IEEE, 178-178.

FERRAZ, G., DE BRITO, J., DE FREITAS, V. & SILVESTRE, J. 2016. State-of-the-art review of building inspection systems. *Journal of performance of constructed facilities,* 30**,** 04016018.

GAVRIKOVA, E., VOLKOVA, I. & BURDA, Y. 2020. Strategic Aspects of Asset Management: An Overview of Current Research. *Sustainability,* 12**,** 5955.

GLOROT, X. & BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010. JMLR Workshop and Conference Proceedings, 249-256.

GORDIAN. 2022. *Surveying Software and Inspections Application with Kykloud* [Online]. Gordian. Available: https://www.gordian.com/products/mobile-surveying-inspections-application-

kykloud/?gclid=CjwKCAiAlrSPBhBaEiwAuLSDUOFsgx_eh4tmFO2TOOKZIIp3PiLbmEOFC073M2_klPvyfRBtV_iFkxoC72IQAvD_BwE [Accessed 23/01/2022 2022].
GRIFFIN, G., HOLUB, A. & PERONA, P. 2007. Caltech-256 object category dataset.
HAN, K. K. & GOLPARVAR-FARD, M. 2015. Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs. *Automation in construction,* 53**,** 44-57.
HE, K., ZHANG, X., REN, S. & SUN, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016a. 770-778.
HE, K., ZHANG, X., REN, S. & SUN, J. Identity mappings in deep residual networks. European conference on computer vision, 2016b. Springer, 630-645.
HOWARD, A. G., ZHU, M., CHEN, B., KALENICHENKO, D., WANG, W., WEYAND, T., ANDREETTO, M. & ADAM, H. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
HU, J., SHEN, L. & SUN, G. Squeeze-and-excitation networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018. 7132-7141.
HUANG, G., LIU, Z., VAN DER MAATEN, L. & WEINBERGER, K. Q. Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. 4700-4708.
HUI, L., PARK, M.-W. & BRILAKIS, I. 2015. Automated brick counting for façade construction progress estimation. *Journal of Computing in Civil Engineering,* 29**,** 04014091.
IANDOLA, F. N., HAN, S., MOSKEWICZ, M. W., ASHRAF, K., DALLY, W. J. & KEUTZER, K. 2016. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. *arXiv preprint arXiv:1602.07360*.
IOFFE, S. & SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. International conference on machine learning, 2015. PMLR, 448-456.
KARPATHY, A. 2016. Cs231n convolutional neural networks for visual recognition. *Neural networks,* 1.
KRIZHEVSKY, A. & HINTON, G. 2009. Learning multiple layers of features from tiny images.
KRIZHEVSKY, A., SUTSKEVER, I. & HINTON, G. E. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems,* 25**,** 1097-1105.
KUZNETSOVA, A., ROM, H., ALLDRIN, N., UIJLINGS, J., KRASIN, I., PONT-TUSET, J., KAMALI, S., POPOV, S., MALLOCI, M. & KOLESNIKOV, A. 2020. The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale.
LIN, M., CHEN, Q. & YAN, S. 2013. Network in network. *arXiv preprint arXiv:1312.4400*.
LU, Q. & LEE, S. 2017. Image-Based Technologies for Constructing As-Is Building Information Models for Existing Buildings. *J COMPUT CIVIL ENG,* 31**,** 4017005.
LUO, H., XIONG, C., FANG, W., LOVE, P. E. D., ZHANG, B. & OUYANG, X. 2018. Convolutional neural networks: Computer vision-based workforce activity assessment in construction. *AUTOMAT CONSTR,* 94**,** 282-289.
MAEDA, H., SEKIMOTO, Y., SETO, T., KASHIYAMA, T. & OMATA, H. 2018. Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone.
MARZOUK, M. & ZAHER, M. 2020. Artificial intelligence exploitation in facility management using deep learning. *Construction innovation,* 20**,** 609-624.
MAYO, G. & KARANJA, P. 2018. Building condition assessments–methods and metrics. *Journal of Facility Management Education and Research,* 2**,** 1-11.
MOBLEY, R. K. 2002. *An introduction to predictive maintenance*, Elsevier.
MOHAMMED, H. J. 2021. The optimal project selection in portfolio management using fuzzy multi-criteria decision-making methodology. *Journal of sustainable finance & investment***,** 1-17.
NISHIKAWA, T., YOSHIDA, J., SUGIYAMA, T. & FUJINO, Y. 2012. Concrete Crack Detection by Multiple Sequential Image Filtering: Concrete crack detection by image processing. *Computer-aided civil and infrastructure engineering,* 27**,** 29-47.

PAULO, P. V., BRANCO, F. & DE BRITO, J. 2014. BuildingsLife: a building management system. *Structure and Infrastructure Engineering,* 10**,** 388-397.

PEZZICA, C., SCHROETER, J., PRIZEMAN, O., JONES, C. & ROSIN, P. 2019. Between images and built form: Automating the recognition of standardised building components using deep learning.

PRAJAPATI, A., BECHTEL, J. & GANESAN, S. 2012. Condition based maintenance: a survey. *Journal of quality in maintenance engineering,* 18**,** 384-400.

QUATTONI, A. & TORRALBA, A. Recognizing indoor scenes. 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009. IEEE, 413-420.

RASHIDI, A., SIGARI, M. H., MAGHIAR, M. & CITRIN, D. 2016. An analogy between various machine-learning techniques for detecting construction materials in digital images. *KSCE Journal of Civil Engineering,* 20**,** 1178-1188.

RICS 2012. New Rules of Measurement. RICS.

RICS. 2021. *RICS NRM: New Rules of Measurement* [Online]. Available: https://www.rics.org/uk/upholding-professional-standards/sector-standards/construction/nrm/ [Accessed January 1 2022].

RUSSAKOVSKY, O., DENG, J., SU, H., KRAUSE, J., SATHEESH, S., MA, S., HUANG, Z., KARPATHY, A., KHOSLA, A. & BERNSTEIN, M. 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision,* 115**,** 211-252.

RUSSELL, B. C., TORRALBA, A., MURPHY, K. P. & FREEMAN, W. T. 2008. LabelMe: a database and web-based tool for image annotation. *International journal of computer vision,* 77**,** 157-173.

SHAHA, M. & PAWAR, M. Transfer learning for image classification. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018. IEEE, 656-660.

SIMONYAN, K. & ZISSERMAN, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

STANDARDIZATION, I. O. F. 2014. ISO 55000: Asset Management-Overview, Principles and Terminology, International Organization for Standardization.

SZEGEDY, C., LIU, W., JIA, Y., SERMANET, P., REED, S., ANGUELOV, D., ERHAN, D., VANHOUCKE, V. & RABINOVICH, A. Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 1-9.

SZEGEDY, C., VANHOUCKE, V., IOFFE, S., SHLENS, J. & WOJNA, Z. Rethinking the inception architecture for computer vision. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 2818-2826.

TOO, E. & TOO, L. 2010. Strategic infrastructure asset management: a conceptual framework to identify capabilities. *Journal of corporate real estate,* 12**,** 196-208.

XIAO, J., HAYS, J., EHINGER, K. A., OLIVA, A. & TORRALBA, A. Sun database: Large-scale scene recognition from abbey to zoo. 2010 IEEE computer society conference on computer vision and pattern recognition, 2010. IEEE, 3485-3492.

YEUM, C. M. & DYKE, S. J. 2015. Vision‐based automated crack detection for bridge inspection. *Computer‐Aided Civil and Infrastructure Engineering,* 30**,** 759-770.

YU, F., SEFF, A., ZHANG, Y., SONG, S., FUNKHOUSER, T. & XIAO, J. 2015. Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv preprint arXiv:1506.03365*.

ZALAMA, E., GÓMEZ‐GARCÍA‐BERMEJO, J., MEDINA, R. & LLAMAS, J. 2014. Road crack detection using visual features extracted by Gabor filters. *Computer‐Aided Civil and Infrastructure Engineering,* 29**,** 342-358.

ZHAN, J., GE, X. J., HUANG, S., ZHAO, L., WONG, J. K. W. & HE, S. X. 2019. Improvement of the inspection-repair process with building information modelling and image classification. *Facilities (Bradford, West Yorkshire, England),* 37**,** 395-414.

ZHOU, B., ZHAO, H., PUIG, X., FIDLER, S., BARRIUSO, A. & TORRALBA, A. Scene parsing through ade20k dataset. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. 633-641.

ZHU, Y., SAPRA, K., REDA, F. A., SHIH, K. J., NEWSAM, S., TAO, A. & CATANZARO, B. Improving semantic segmentation via video propagation and label relaxation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 8856-8865.

ZHU, Z., GERMAN, S. & BRILAKIS, I. 2010. Detection of large-scale concrete columns for automated bridge inspection. *Automation in construction,* 19**,** 1047-1055.

APPENDIX

Appendix 1 The Summarization Table of Different General-purpose and AEC/FM Built Assets related Image Datasets

Dataset Name	Dataset Type	Brief Description	Sample Categories	Preprocessing	Instances	Format	Default Task	Created (or updated)	Reference	Creator
Caltech-101	**General-purpose**	Pictures of objects belonging to 101 categories. About 40 to 800 images per category. Most categories have about 50 images.	Pigeon, airplanes, helicopter, barrel, lobster, brain, etc.	Image labelled	9,146	Images	Classification, object recognition.	2003	(Fei-Fei et al., 2004)	F. Li et al.
Caltech-256	**General-purpose**	Consists of a large dataset of images for object classification. There are 30,607 images in this dataset spanning 257 object categories. Object categories are extremely diverse, ranging from grasshopper to tuning fork.	Sports (e.g., golf-ball, soccer-ball, tennis-ball, etc.), fun (e.g., superman, cartman, teddy bear, etc.), electronics (e.g., cd, iPod, washing-machine, etc.), structure (e.g., skyscraper, pyramid, lighthouse, etc.), etc.	Images categorized and hand-sorted.	30,607	Images	Classification, object detection	2007	(Griffin et al., 2007)	G. Griffin et al.
LabelMe	**General-purpose**	Contains more than 1,000 fully annotated images and around 2,000 partially annotated images. Including partially annotated images allows algorithms to show if they are able to benefit from additional partially labelled images.	Images, Objects, Cars, Person, Building (e.g., house, paper cup, chair, table, chair, etc.), Road (e.g. car), Sidewalk, Sky, Tree, etc.	Image labelled, training set splits are created.	96,115	Images	Classification, object detection	2008	(Russell et al., 2008)	B. Russell et al.
CIFAR-10 Dataset	**General-purpose**	Consists of 60,000 32×32 colour images in 10 classes, with 6,000 images per class. There are 50,000 training images and 10,000 test images.	Airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck, etc.	Image labelled, training set splits created.	60,000	Images	Classification	2009	(Krizhevsky and Hinton, 2009, Krizhevsky et al., 2012)	A. Krizhevsky et al.
CIFAR-100 Dataset	**General-purpose**	Has 100 classes containing 600 images each. There are 500 training images and 100 testing images per class. The 100 classes in the CIFAR-100 are grouped into 20 superclasses. Each image comes with a "fine" label (the class to which it belongs) and a "coarse" label (the superclass to which it belongs).	Aquatic mammals (e.g., beaver, dolphin, seal, etc.), fish (e.g., aquarium fish, ray, shark, etc.), household electrical devices (e.g., clock, computer keyboard, lamp, etc.), Household furniture (bed, chair, couch, table, etc.), etc.	Image labelled, training set splits created.	60,000	Images	Classification	2009	(Krizhevsky and Hinton, 2009, Krizhevsky et al., 2012)	A. Krizhevsky et al.
PASCAL VOC Dataset	**General-purpose**	Contains a large number of images for classification tasks. In the year 2012, it contains 20 classes. The train and validation datasets has 11,530 images containing 27,450 regions of interest (ROI) annotated objects and 6,929 segmentations.	Person (person), animal (e.g., bird, cat, cow, etc.), vehicle (e.g., aeroplane, bicycle, boat, etc.), indoor (e.g., bottle, chair, dining table, etc.), etc.	Image labelled, bounding box included	500,000	Images	Classification, object detection	2012	(Everingham et al., 2010)	M. Everingham et al.
SUN Database	**General-purpose**	Consists of a very large sized scene and object recognition database, where UNderstanding (SUN) database contains 899 categories and 130,519 images.	Scene categories: abbey, badminton court indoor, cargo container interior, utility room, dining room, garage outdoor, jail indoor, etc. Object categories: ceiling, window,	Places and objects are labelled. Objects are segmented.	131,067	Images	Object recognition, scene recognition	2014	(Xiao et al., 2010)	J. Xiao et al.

			cabinet, sink, chair, bottle, circus, table, football, food platter, etc.							
ImageNet	**General-purpose**	ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which hundreds and thousands of images depict each node of the hierarchy. Currently, it has an average of over five hundred images per node.	plant, flora, plant life; geological formation, formation; natural object; sport, athletics; artifact, artefact (e.g., instrumentality & instrumentation, structure & construction, paving & pavement, etc.); fungus; person, individual, someone; animal, animate being; Misc; etc.	Image labelled, bounding boxes, descriptive words, SIFT features	14,197,122	Images	Object recognition, scene recognition	2014	(Deng et al., 2009)	J. Deng et al.
Microsoft Common Objects in Context (COCO)	**General-purpose**	COCO is large-scale object detection, segmentation, and captioning dataset with more than 330,000 images, 1.5 million object instances, 80 object categories, 91 stuff categories, 5 captions per image.	Person, airplane, bird, horse, skis, banana, apple, chair, couch, sink, hair drier, etc.	Object highlighted, labelled, and classified.	1,500,000	Images	Object recognition	2015	(Lin et al., 2013)	T. Lin et al.
LSUN	**General-purpose**	Contains around one million labelled images for each of 10 scene categories and 20 object categories.	Airplane, bicycle, bird, bus, sofa, train, etc.	Image labelled	10,000,000	Images	Classification	2016	(Yu et al., 2015)	Fisher Yu et al.
Linnaeus 5 dataset	**General-purpose**	Contains images from 5 object classes. Images are 256×256 pixels. 1,200 training images, 400 test images per class.	Berry, bird, dog, flower, Other, etc.	Image labelled, training set splits are created.	8,000	Images	Classification	2017	(Chaladze and Kalatozishvili, 2017)	Chaladze, G. Kalatozishvili L
ADE20K	**General-purpose**	Contains more than 20K scene-centric images exhaustively annotated with objects and object parts. In addition, there are a total of 150 semantic categories included for evaluation.	Sky, road, grass, person, card, bed.	Object highlighted, labelled, and classified.	22,210	Images	Classification, object recognition, scene recognition	2017	(Zhou et al., 2017)	B. Zhou et al.
CINIC-10 Dataset	**General-purpose**	A unified contribution of CIFAR-10 and Imagenet with 10 classes, and 3 splits. CINIC-10 is a drop-in replacement for CIFAR-10. It is compiled as a benchmarking dataset because CIFAR-10 can be too small/too easy, and ImageNet is often too large/difficult. ImageNet32 and ImageNet64 are smaller than ImageNet but even more difficult. CINIC-10 fills this benchmarking gap.	The subcategories of CIFAR-10 and ImageNet	Image labelled, training set splits created.	270,000	Images	Classification	2018	(Darlow et al., 2018)	Luke N. Darlow et al.
Open Images	**General-purpose**	The images included are very diverse and often contain complex scenes with several objects (8.4 per image on average). It contains image-level labels annotations, object bounding boxes, object segmentation, visual relationships, and localized narratives.	Building (e.g., sports equipment, music equipment, camera, billard table, bottle opener), food, vehicle, animal, furniture, tool, etc.	Image labelled, Bounding boxes	9,178,275	Images	Classification, Object recognition	2020	(Kuznetsova et al., 2020)	A. Kuznetsova et al.
Indoor Scene Recognition	**AECFM related**	Contains 67 Indoor categories with a total of 15620 images. The number of images varies across categories, but there are at least 100 images per category. All images are in jpg	Store (e.g., bakery, grocery store, clothing store, bookstore, etc.), Home (bedroom, nursery, closet, pantry, bathroom, etc.), public spaces (e.g., prison cell, library, church, museum, etc.), leisure (buffet, fast food, restaurant, bar, gym, etc.), working	Image labelled	15,620	Images	Classification	2009	(Quattoni and Torralba, 2009)	A. Quattoni, A.Torralba

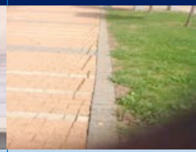
		format. The images provided here are for research purposes only.	place (e.g., hospital room, restaurant kitchen, classroom, laboratory, etc.), etc.							
CML	**AECFM related**	CML is a new database containing 22 typical construction materials with more than 150 images per category is assembled	asphalt, brick, cement-granular, cement-smooth, concrete-cast, concrete-precast, foliage, formwork, grass, gravel, marble, metal-grills, paving, soil-compact, soil-vegetation, soil-loose, soil-mulch, stone-granular, stone-limestone, wood.	Image labelled	4,400	Images	Classification	2015	(Han and Golparvar-Fard, 2015)	K. Han, M. Golparvar-Fard
MINC	**AECFM related**	MINC is a 3-million-images diverse dataset that has 23 different material related categories with more examples in less common categories. MINC draws data from both Flickr images, which includes many “regular” scenes and Houzz images from professional photographers of staged interiors.	Brick, carpet, ceramic, fabric, foliage, food, glass, hair, leather, metal, mirror, other, painted, paper, plastic, pol. stone, skin, sky, stone, tile, wallpaper, water, wood.	Object highlighted, labelled, and classified into 23 object types.	3,000,000	Images	Classification, Object detection	2015	(Bell et al., 2015)	S. Bell et al.
MCIndoor 20000	**AECFM related**	MCIndoor20000 is a 20,000-image dataset contains digital images from three different indoor object categories.	Doors, stairs, and hospital signs.	Image labelled	20,000	Images	Classification	2018	(Bashiri et al., 2018)	Bashiri et al.
Carnegie library dataset	**AECFM related**	Carnegie library dataset is a 13,000-image dataset contain image record of Carnegie library building stock in the UK,	Turret, tower, grille exterior, grille interior.	Image labelled	13,000	Images	Classification, Object detection	2019	(Pezzica et al., 2019)	Pezzica et al.
BIM knowledge repository	**AECFM related**	BIM knowledge repository contains 3D model, object information and activity histories (for the inspection, assignment and repair) are saved in a Navisworks file. Other information, including images taken by inspectors or repairmen and user comments on activities, are stored in image files (.jpg or .png) or plain text files (.txt).	Objects (e.g., refrigerator, microwave, coffee machine, sink, chair and monitor) in the physical environment	Image labelled	60	Images	Classification, object detection	2019	(Zhan et al., 2019)	J. Zhan et al.
FM200	**AECFM related**	A set of images for fire protection systems, with 3 Classes, in the proposed classification, namely, FM200, fire extinguisher and fire hose cabinet has been downloaded. The downloaded images are having different sizes.	Fire extinguisher and fire hose cabinet	Image labelled, training set splits created.	81	Images	Classification, Object recognition	2020	(Marzouk and Zaher, 2020)	Mohamed Marzouk et al.

Appendix 2 The Summarization Table of the Detailed Project Portfolio Information

Project Portfolio Name	Project Name	GIFA - Portfolio	Location	Region	Project Type	Description of Works	Lifecycle expenditure - Portfolio Total	Delivery Method	Number of Images	Start Year - Portfolio	Number of Image Categories
ImageA	ImageAA	45,210	Birmingham, UK	West Midlands England, UK	Education	Defect Survey	£ 4,200,000.00	Subcontracted	130	2017	10
ImageB	ImageBA	5,901	Cumbria Lift, UK	North West England, UK	Health	Condition Survey	£ 2,300,000.00	Subcontracted	465	2016	51
ImageC	ImageCA	83,337	Blackburn, UK	North West England, UK	Health	Defect Survey	£ 72,000,000.00	Subcontracted	1,836	2016	108
ImageD	ImageDA	71,708	London, UK	London Region, UK	Health	Condition Survey	£ 35,000,000.00	Subcontracted	254	2016	36
ImageE	ImageEA	-	London, UK	London Region, UK	Housing	Condition Survey	£ 2,400,000.00	Subcontracted	206	2015	26
ImageE	ImageEB	-	London, UK	London Region, UK	Housing	Condition Survey	£ 2,400,000.00	Subcontracted	11	2015	2
ImageF	ImageFA	35,671	Swindon, UK	South West England, UK	Education	Condition Survey	£ 17,000,000.00	Subcontracted	175	2016	33
ImageF	ImageFB	35,671	Swindon, UK	South West England, UK	Education	Condition Survey	£ 17,000,000.00	Subcontracted	348	2016	101
ImageF	ImageFC	35,671	Swindon, UK	South West England, UK	Education	Condition Survey	£ 17,000,000.00	Subcontracted	431	2016	89
ImageF	ImageFD	35,671	Swindon, UK	South West England, UK	Education	Condition Survey	£ 17,000,000.00	Subcontracted	64	2016	41
ImageF	ImageFE	35,671	Swindon, UK	South West England, UK	Education	Condition Survey	£ 17,000,000.00	Subcontracted	133	2016	72
ImageF	ImageFF	35,671	Swindon, UK	South West England, UK	Education	Condition Survey	£ 17,000,000.00	Subcontracted	51	2016	30
ImageF	ImageFG	35,671	Swindon, UK	South West England, UK	Education	Condition Survey	£ 17,000,000.00	Subcontracted	30	2016	16

ImageG	ImageGA	-	London, UK	London Region, UK	Education	Condition Survey	£	3,500,000.00	Subcontracted	1,088	2011	83
ImageG	ImageGB	-	London, UK	London Region, UK	Education	Condition Survey	£	3,500,000.00	Subcontracted	876	2011	96
ImageG	ImageGC	-	London, UK	London Region, UK	Education	Condition Survey	£	3,500,000.00	Subcontracted	216	2011	62
ImageG	ImageGD	-	London, UK	London Region, UK	Education	Condition Survey	£	3,500,000.00	Subcontracted	118	2011	48
ImageH	ImageHA	8,615	Kingston, UK	London Region, UK	Health	Condition Survey	£	7,900,000.00	Subcontracted	469	2017	92
ImageI	ImageIA	35,912	North Staffs, UK	West Midlands England, UK	Health	Condition Survey	£	100,000,000.00	Subcontracted	195	2017	58
ImageJ	ImageJA	-	Pembury, UK	South East England, UK	Health	Condition Survey	£	41,000,000.00	Subcontracted	1,977	2018	319
ImageJ	ImageJB	-	Pembury, UK	South East England, UK	Health	Condition Survey	£	41,000,000.00	Subcontracted	3,557	2018	123

Appendix 3 The Visualization Table of Five Randomly Selected Test Images with Their Prediction for Three Selected Models against Different Training Dataset Settings

Category Name:		CCTV Camera	Internal Doors Frames & architraves	Fire Suppression Systems Sprinkler Pump Controls panels inverter drives	Roads Paths and Paving Concrete	Commercial 4 way hobs
		05.12.02.20	02.08.01.25	05.11.02.03	08.02.01.06	05.02.02.06
Demo Images						
MobileNet1.0	Both On-site and Online Photos	Internal Doors Frames & architraves 02.08.01.25 24.9%	Internal Doors Frames & architraves 02.08.01.25 89.3%	Distribution Boards 05.08.02.02 85.0%	Roads Paths and Paving Concrete 08.02.01.06 62.5%	Commercial Deep fat fryers 05.02.02.04 23.5%
		Internal Door's Ironmongery 02.08.01.02 24.2%	Internal Timber Single Door 02.08.01.11 10.1%	Fire Suppression Systems Sprinkler Pump Co 05.11.02.03 4.5%	External Doors Aluminium 02.06.02.01 9.6%	Commercial Hot cupboard units 05.02.02.15 18.4%
		Internal Timber Double door 02.08.01.13 21.7%	Internal Door's Ironmongery 02.08.01.02 0.5%	Electric Mains and Sub mains Distribution In 05.08.01.01 2.0%	Walls and Screens Brick 08.04.02.01 6.2%	Commercial Gas cooker units 05.02.02.01 16.1%
		Internal Timber Single Door 02.08.01.11 9.4%	Internal Timber Double door 02.08.01.13 0.1%	Motor Control Cubicles 05.08.01.12 1.5%	Natural stone External Walls 02.05.01.01 2.7%	Commercial Main Dishwasher 05.02.02.05 14.5%
		Window's Ironmongery 02.06.01.16 1.8%	External Doors Aluminium 02.06.02.01 0.0%	Low Voltage Main Switchboard 05.08.01.11 1.4%	Tarmacadam Paving 08.02.01.26 2.2%	Commercial Kitchen Worktops 05.02.02.25 14.0%
	On-site Photos Only	Access Control Headend 05.12.02.03 68.6%	Internal Doors Frames & architraves 02.08.01.25 99.9%	Fire Suppression Systems Sprinkler Pump Co 05.11.02.03 63.4%	Roads Paths and Paving Concrete 08.02.01.06 96.3%	Commercial Deep fat fryers 05.02.02.04 68.9%
		Fire fighting Systems Dry Risers 05.11.01.01 6.2%	Internal Timber Single Door 02.08.01.11 0.1%	Distribution Boards 05.08.02.02 12.7%	Walls and Screens Brick 08.04.02.01 1.0%	Commercial Fridge freezer units Combined 05.02.02.11 11.5%
		Vandal proof Sealed Fluorescent luminaire 05.08.03.05 5.4%	Internal Door's Ironmongery 02.08.01.02 0.0%	Pump Control panels 05.04.02.13 10.4%	Natural stone External Walls 02.05.01.01 0.8%	Commercial Gas cooker units 05.02.02.01 7.7%
		Communication Systems Telecoms Data Eq 05.12.01.02 5.0%	Internal Timber Double door 02.08.01.13 0.0%	Wiring and cables from local distribution bo 05.08.03.15 4.7%	Roads Paths and Paving Cobbles 08.02.01.29 0.2%	Commercial Cold fridge counters 05.02.02.13 3.9%
		Passenger Goods Lift Electric Traction 05.10.01.01 1.9%	Leaf and half Timber door 02.08.01.12 0.0%	Commercial Photovoltaic Electricity Genera 05.08.05.04 2.6%	Block wall 02.05.01.08 0.2%	Commercial Hot cupboard units 05.02.02.15 3.5%
ResNet152 v1	Both On-site and Online Photos	Internal Doors Sliding concertina 02.08.01.08 49.1%	Internal Doors Frames & architraves 02.08.01.25 94.2%	Distribution Boards 05.08.02.02 51.0%	Tarmacadam Paving 08.02.01.26 63.1%	Commercial Gas cooker units 05.02.02.01 32.0%
		Hot water storage cylinder 05.04.03.03 21.9%	Internal Timber Single Door 02.08.01.11 4.9%	Pump Control panels 05.04.02.13 18.3%	Roads Paths and Paving Concrete 08.02.01.06 32.9%	Commercial Combi Oven units 05.02.02.16 17.6%
		lightning protection down conductors 05.11.03.01 6.3%	Internal Door's Ironmongery 02.08.01.02 0.8%	Fire Suppression Systems Sprinkler Pump Co 05.11.02.03 18.3%	Roads Paths and Paving Concrete kerbs 08.02.01.08 2.2%	kitchen Hot Plate unit 05.02.02.30 14.6%
		Internal Doors Frames & architraves 02.08.01.25 3.8%	Internal Doors Paint to timber 02.08.01.05 0.0%	Fire Alarm Control Panel 05.12.01.03 3.3%	External Doors Aluminium 02.06.02.01 0.8%	Commercial Griddle units 05.02.02.07 12.7%
		Cold water distribution Steel Storage Tank 05.04.02.01 3.3%	Internal Timber Double door 02.08.01.13 0.0%	Fire Alarm Wiring system 05.12.01.07 2.1%	Walls and Screens Brick 08.04.02.01 0.2%	Commercial Main Dishwasher 05.02.02.05 7.3%
	On-site Photos Only	Access Control Headend 05.12.02.03 93.2%	Internal Doors Frames & architraves 02.08.01.25 100.0%	Fire Suppression Systems Sprinkler Pump Co 05.11.02.03 66.9%	Roads Paths and Paving Concrete 08.02.01.06 37.0%	Commercial Deep fat fryers 05.02.02.04 98.0%
		BMS Local Control Panels 05.12.03.04 3.9%	Internal Timber Single Door 02.08.01.11 0.0%	Distribution Boards 05.08.02.02 26.6%	External Doors Aluminium 02.06.02.01 21.8%	Commercial Gas cooker units 05.02.02.01 1.0%
		Fire fighting Systems Dry Risers 05.11.01.01 1.2%	Internal Door's Ironmongery 02.08.01.02 0.0%	Pump Control panels 05.04.02.13 1.7%	Roads Paths and Paving Concrete kerbs 08.02.01.08 7.8%	Commercial Hot cupboard units 05.02.02.15 0.2%
		Communication Systems Telecoms Data Eq 05.12.01.02 0.5%	Internal Timber Double door 02.08.01.13 0.0%	Motor Control Cubicles 05.08.01.12 1.4%	Timber edging 08.02.01.27 7.3%	Waste disposal units 05.02.02.09 0.2%
		Vandal proof Sealed Fluorescent luminaire 05.08.03.05 0.3%	Paint plaster wall 03.01.01.06 0.0%	Auto Transfer Switch 05.08.01.08 1.0%	Natural stone External Walls 02.05.01.01 5.9%	Commercial Cold fridge counters 05.02.02.13 0.2%
ResNet18 v1	Both On-site and Online Photos	Internal Door's Ironmongery 02.08.01.02 56.8%	Internal Doors Frames & architraves 02.08.01.25 81.9%	Distribution Boards 05.08.02.02 78.9%	Roads Paths and Paving Concrete 08.02.01.06 45.5%	Commercial Gas cooker units 05.02.02.01 17.4%
		Internal Doors Frames & architraves 02.08.01.25 23.8%	Internal Timber Single Door 02.08.01.11 16.2%	Pump Control panels 05.04.02.13 6.1%	Natural stone External Walls 02.05.01.01 30.4%	Commercial Deep fat fryers 05.02.02.04 16.7%
		lightning protection down conductors 05.11.03.01 2.5%	Internal Door's Ironmongery 02.08.01.02 1.8%	Fire Suppression Systems Sprinkler Pump Co 05.11.02.03 4.6%	External Doors Aluminium 02.06.02.01 6.2%	Commercial Griddle units 05.02.02.07 16.0%
		Metal post and rail 08.04.01.12 2.0%	Internal Timber Double door 02.08.01.13 0.0%	Fire fighting Systems Dry Risers 05.11.01.01 3.2%	Tarmacadam Paving 08.02.01.26 4.6%	Commercial Hot cupboard units 05.02.02.15 12.9%
		Grab rails 04.01.03.21 1.6%	Internal Doors Paint to timber 02.08.01.05 0.0%	Low Voltage Main Switchboard 05.08.01.11 1.7%	Walls and Screens Brick 08.04.02.01 2.8%	Commercial Grills 05.02.02.08 12.6%
	On-site Photos Only	Internal Timber Double door 02.08.01.13 62.4%	Internal Doors Frames & architraves 02.08.01.25 99.5%	Fire Suppression Systems Sprinkler Pump Co 05.11.02.03 78.7%	Roads Paths and Paving Concrete 08.02.01.06 36.4%	Commercial Main Dishwasher 05.02.02.05 23.4%
		BMS Local Control Panels 05.12.03.04 8.7%	Internal Timber Single Door 02.08.01.11 0.3%	Distribution Boards 05.08.02.02 12.9%	External Doors Aluminium 02.06.02.01 15.4%	Commercial Gas cooker units 05.02.02.01 17.2%
		Communication Systems Telecoms Data Eq 05.12.01.02 2.0%	Internal Door's Ironmongery 02.08.01.02 0.1%	UPS System 05.08.02.04 2.6%	Roof stone coated 02.03.02.34 7.3%	Commercial Griddle units 05.02.02.07 13.7%
		Internal Timber Single Door 02.08.01.11 1.8%	Boiler Plant Gas condensing 05.05.01.03 0.1%	Cold water supply Booster Pump set 05.04.02.12 1.4%	Roads Paths and Paving Cobbles 08.02.01.29 4.4%	Commercial Deep fat fryers 05.02.02.04 12.2%
		Medical gas Pipe Installation 05.13.01.01 1.6%	Internal Timber Double door 02.08.01.13 0.0%	Pump Control panels 05.04.02.13 0.8%	Roads Paths and Paving Concrete kerbs 08.02.01.08 4.1%	Waste disposal units 05.02.02.09 4.8%