Artificial intelligence applications for proactive environmental monitoring and asset management

Jun Kang Chow[1], **Pin Siang Tan**[1], **Kuan-fu Liu**[1], **Xin Mao**[1], **Zhaoyu Su**[1], **Ghee Leng Ooi**[1], **Ye Hur Cheong**[1], **Mei Ling Leung**[1], **Jimmy Wu**[1], **Hok Man Chan**[2], **Lai Yuk Yip**[2], **Ka Chun Chow**[2] **and Yu-Hsing Wang**[2]

[1] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, People's Republic of China

[2] Drainage Services Department, the HKSAR Government, Hong Kong, People's Republic of China

ABSTRACT

Two research studies have been implemented to explore the potential of applying artificial intelligence (AI) technologies in works projects and maintenance work of the Drainage Services Department (DSD) for enhancing the efficiency related to environmental monitoring and structural inspection, referred to as the AIEIA and AIBIM projects, respectively. In the AIEIA project, AI technologies were explored to assist observing bird behaviour that would likely be influenced by nearby DSD construction projects. A domain randomisation-enhanced model was built to detect great egrets and little egrets at Penfold Park, Hong Kong, achieving a mean average precision of 87.65%. The detection result was used to analyse the Penfold Park egretry behaviour. In the AIBIM project, AI technologies were used to facilitate the condition assessment of concrete defects in sewage treatment facilities. A classifier was developed with supervised learning for concrete defect detection, attaining recalls of 86.2% and 89.9% for the cracking and spalling classes. Another concrete defect anomaly detector was built using unsupervised learning, achieving balanced results with F_2 measures of 85.2% and 76.0% for the cracking and spalling classes. The two research studies render valuable experience for the DSD to integrate AI-enabled analytics into future work to continuously improve the drainage services in Hong Kong.

KEYWORDS Deep learning; domain randomisation; egretry behaviour; convolutional layers; concrete defects; supervised learning; unsupervised learning

CONTACT Yu-Hsing Wang ceyhwang@ust.hk

Received 9 November 2021

1. Introduction

The Drainage Services Department (DSD) envisions providing world-class wastewater and stormwater drainage services to enable the sustainable development of Hong Kong. To support this vision, the DSD keeps abreast of industry developments and explores ways to improve its day-to-day work through the application of innovative technologies. In this connection, the DSD has taken the opportunity to collaborate with The Hong Kong University of Science and Technology (HKUST) to explore integrating state-of-the-art technologies, particularly artificial intelligence (AI). As shown in Figure 1, this paper highlights and discusses two collaborative research projects, named (i) Big Data Artificial Intelligence-based Environmental Impact Assessment of Relocation of Sha Tin Sewage Treatment Works to Caverns (abbreviated as AIEIA) and (ii) Development of Smart System for Asset Management of Structural Concrete – Artificial Intelligence-based Building Information Modelling (abbreviated as AIBIM), that explore the potential of deploying AI technology, especially the deep learning-based frameworks developed by Chow et al. (2020a and 2020b) and Mao et al. (2021), in environmental monitoring and routine structural inspection works of the DSD.

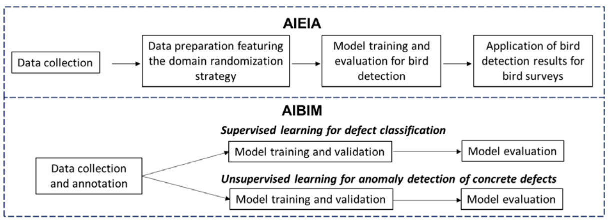


Figure 1. An overview of the two collaborative research projects, AIEIA and AIBIM.

2. AIEIA project

2.1. Project background

The relocation of the Sha Tin Sewage Treatment Works (STSTW) to caverns in Nui Po Shan aims at releasing the existing site for housing and other beneficial uses and improving the environment of the existing site and its surroundings. Egrets are commonly seen near the existing STSTW as they forage in the adjacent Shing Mun River and Sha Tin Hoi, and inhabit the Penfold Park egretry, which is approximately 750 metres away from the existing STSTW. Glare and noise generated from construction, decommissioning and demolition works at the existing STSTW may affect the various activities of the egrets.

In the current practice in Hong Kong, the monitoring of birds and their habitats is mainly performed by manual observation at a specified time and date. This method requires qualified ecologists to be present on site to exercise their expert judgement, and by spreading their efforts to cover multiple spots of interest over the survey site, they very often can only afford a limited period of time at each spot. These manual observation surveys are thus resource-demanding to conduct, and the survey results are only available during the site visit period by the ecologists, as continuous monitoring is practically impossible. Moreover, environmental factors like adverse weather conditions or dense vegetation may influence the observations by human eyes. Hence, the use of AI technology to augment the traditional manual methods was conceived and the feasibility of applying AI in environmental monitoring was studied through a pilot trial of observing the Penfold Park egretry.

This section on the AIEIA project is organised as follows. Data collection at the study site is first briefly described. Then, details of data preparation for the development of deep learning models for bird detection, particularly the domain randomisation strategy used to enhance the learning of fine-grained bird features, are presented. The deep learning models were trained and evaluated, and the detection results were used to analyse the egretry behaviour at the study site.

2.2. Data collection

A tailor-made Green AI Camera system (Mao et al., 2021) was invented to implement automated data collection outdoors. The novel Green AI Camera system comprises six 4K-resolution cameras configured to record videos at a rate of 25 frames per second with High Efficiency Video Coding. The volume of videos is large, e.g., a 20-minute video recording is about 20 gigabytes; therefore, multiple Solid-State Drives (SSDs) are used to provide sufficiently large local data storage, and the SSDs with full memory can be simply unplugged and replaced with fresh ones for continuous storage of the acquired videos. Penfold Park, Hong Kong was chosen as the study site, where great egrets (*Ardea alba*) and little egrets (*Egretta garzetta*) are the dominant bird species, with other birds including black-crowned night herons (*Nycticorax nycticorax*), Chinese pond herons (*Ardeola bacchus*) and eastern cattle egrets (*Bubulcus ibis*) inhabiting the area. Therefore, great egrets and little egrets were chosen as the main monitoring targets. The Green AI Camera system was installed (see Figure 2 for the location) to acquire videos of birds inhabiting trees at the centre of the pond, from 04:00 to 20:00 throughout the monitoring period.

Figure 2. Study site of the AIEIA project. The Green AI Camera system was installed at Penfold Park, Hong Kong to monitor the egretry behaviour.

2.3. Data preparation featuring the domain randomisation strategy

The video frames collected between 13 May and 13 September 2019 were annotated, i.e., bounding boxes were drawn with labels assigned for the observed birds, to facilitate the supervised learning of deep learning models. The observed birds were grouped into three classes, i.e., great egrets, little egrets and other birds. A total of 1,110 images were annotated, and the labelled images were split into 900, 100 and 110 images as the training, validation and testing sets. However, such a small volume of labelled data was unfavourable for satisfactory model training.

The data volume could be enlarged by incorporating images of open-source databases. However, the existing image databases for Aves detection (e.g., CUB-200-211, iNaturalist) were found to be inappropriate for use in the study, after noting that the viewing perspective and representation of birds were different. For instance, bird images in the existing databases are usually close-up images, whereas birds found in the video frames collected at the study site were relatively small in size (represented by a limited number of pixels; see the sample video frame in Figure 2). In addition, deep learning models might encounter difficulties in efficiently learning the fine-grained bird features from a small volume of image data to classify bird species.

The domain randomisation strategy was adopted to resolve the aforementioned challenges. Domain randomisation uses virtual objects and/or environments to augment the data volume and variability (Peng et al., 2018), and has been found to be effective in different applications, e.g., collision avoidance (Sadeghi et al., 2016) and robotic control for grasping specific objects (Tobin et al., 2017). In this study, two virtual birds, i.e., a great egret and a little egret (see Figures 3(a) and 3(b)), were created using the open-source 3D graphics toolset Blender (v2.79). Then, virtual birds were rendered onto backgrounds taken at the study site to create synthetic images (see Figure 3c for examples). A wide range of variations in bird pose and environmental conditions (e.g., the absence/presence of sunlight and obstruction by tree branches) were applied to guide the model to focus on learning fine-grained bird features, such as head and neck. The bird size distribution was set similar to the egret size found in the recorded video. More details of deploying domain randomisation in creating synthetic egrets can be found in Mao et al. (2021).

(a) Virtual great egret (b) Virtual little egret

(c) Syntheitc images

Figure 3. Domain randomisation strategy (Mao et al., 2021) was applied to enhance model training for bird detection: (a) virtual great egret; (b) virtual little egret; and (c) examples of synthetic images.

2.4. Model training of deep learning models for bird detection

The bird detection was considered as an object detection task, i.e., plotting a bounding box on an image for the identified object and assigning an appropriate label. A state-of-the-art deep learning model that primarily uses convolutional layers, i.e., Faster R-CNN (Ren et al., 2017), was selected (see Figure 4 for the network architecture of Faster R-CNN). Convolutional layers exhibit the excellent properties of sparse interaction, translation invariance and parameter sharing, where filters with a size smaller than inputs can be efficiently used to extract representative features of pixelated data, e.g., images (Bengio, 2009). Faster R-CNN is a two-stage detector, where in the first stage, feature maps are extracted with the backbone made up of multiple convolutional layers from the input image, and potential region of interest pooling containing target objects are proposed, and in the second stage, the bounding box and the class of the objects are determined (Ren et al., 2017). In the study, the Feature Pyramid Network (Lin et al., 2017) was used in the region proposal network of Faster R-CNN to efficiently fuse the information of multi-scale feature maps to augment the detection of minute birds found in the video frames.

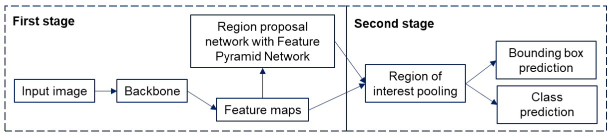


Figure 4. Network architecture of Faster R-CNN used in the study for bird detection.

The model training was implemented using two GeForce RTX 2080 Ti GPUs, with the following training configuration: a batch size of four; the categorial cross-entropy as the loss function; and the stochastic gradient descent with a weight decay of 0.0001, a momentum of 0.9 and a learning rate of 0.0005 as the optimiser. Using the domain randomisation strategy, the Faster R-CNN was first trained with the synthetic images, then fine-tuned with the labelled training set (900 images with real birds recorded at the study site). The model with the best performance compared to the validation set was selected for performance evaluation.

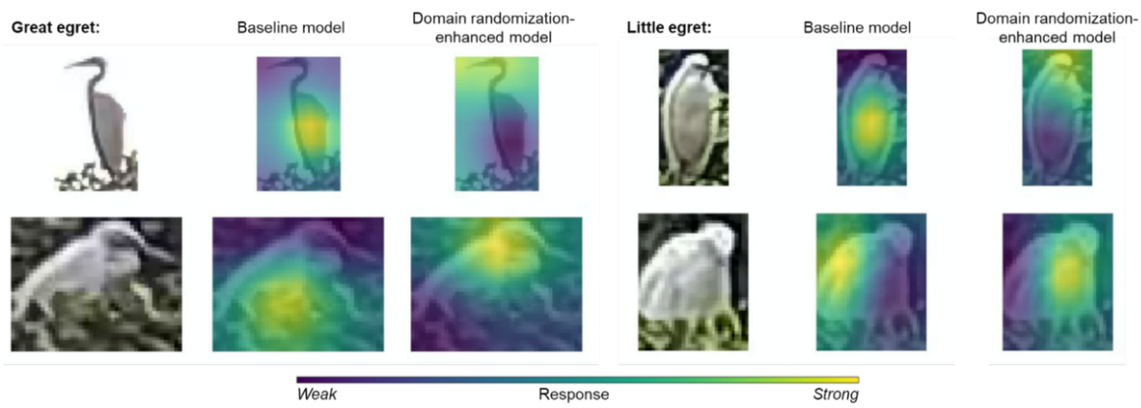


Figure 5. Example of the prediction generated by the domain randomisation-enhanced model, with the red and blue bounding boxes representing the great egrets and little egrets, respectively (updated from Mao et al., 2021).

2.5. Results and discussion

The mean average precision (mAP) at the intersection over union of 0.5 was used to evaluate the performance of the domain randomisation-enhanced model compared to the test set, i.e., 110 real images of birds. The mAP was computed as follows:

$$\mathrm{mAP} = \frac{1}{11}\sum_{\mathrm{Recall}_i \in \{0,0.1,\ldots,0.9,1\}} \mathrm{Precision}(\mathrm{Recall}_i) \quad (1)$$

As shown in Figure 5, the domain randomisation-enhanced model can detect egrets found at different locations, e.g., inhabiting the tops of trees and perching on branches covered partially by leaves, achieving a mAP of 87.65%.

For further evaluation, a baseline model, i.e., Faster R-CNN trained with real images only, was built. Figure 6 presents the perception used by the two deep learning models for detecting egrets. Compared to the baseline model, the domain randomisation-enhanced model tends to focus on the subtle features of the neck and head, consistent with how bird experts differentiate bird species. More details of the quantitative evaluation are discussed in Mao et al. (2021).

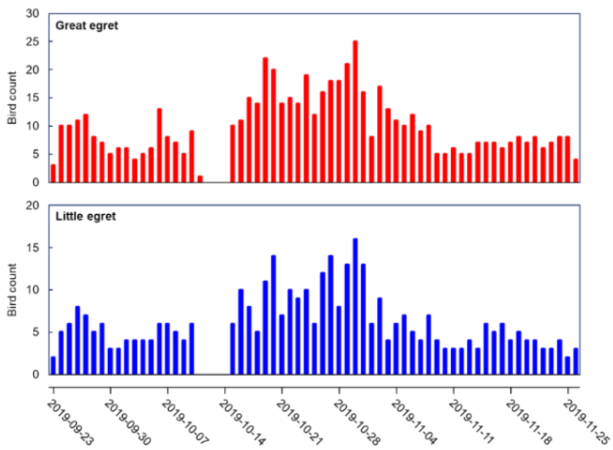
Figure 6. Heatmaps representing the perception made by the deep learning models for bird detection (updated from Mao et al., 2021).

Following model evaluation, the video collected between 23 September and 26 November 2019 (2-month data) was used to study the behaviour of the Penfold Park egrets. Video frames were first analysed by the domain randomisation-enhanced model to detect bird species, i.e., great egrets, little egrets and other birds, and then the detection results were used for different applications of bird surveys.

One of the applications automates daily bird counting. Figure 7(a) presents the daily counts of egrets during the study period. The daily counts are computed based on the highest number of birds detected in a single frame per day and assumed to be the number of birds inhabiting the study site. An increase in the great egret and little egret counts was observed between the end of October and the beginning of November, which is plausibly due to the influx of the migratory population (Carey, 2001; The Hong Kong Bird Watching Society, 2018).

Another application is studying the daily schedule of egrets. Figure 7(b) presents the departure and return time of the egrets. For each individual day, the departure and return times of the egrets were estimated by obtaining the time points with the highest average value in the morning and evening from the backward one-hour moving average of the bird counts for all time points. The departure times of the great egrets and little egrets were found to be similar, whereas the great egrets returned slightly earlier than the little egrets on most of the days during the study period. The analysed daily schedule of egrets can then be used to further investigate the prey availability in different foraging habitat options. For example, in Hong Kong, great egrets commonly search for food at mudflats and little egrets mostly forage at commercial fishponds (Choi et al., 2007; Young, 1998).

(a)



(b)

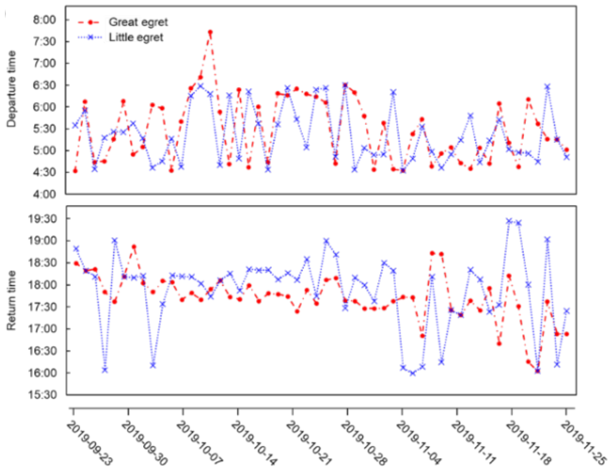


Figure 7. Analysis of the egretry behaviour at the study site from a temporal perspective: (a) the egret counts; and (b) the departure and return time of egrets during the study period (updated from Mao et al., 2021).

The detection results of the deep learning models could also be applied to investigate the spatial distribution of egrets at the study site, as shown in Figure 8. A heatmap was created using the random field theory, where the intensity of the red and blue colours reflects the great egret and little egret counts at a specific location, respectively. Along the horizontal level, the great egrets and little egrets were observed to inhabit the centre regions of trees. On the other hand, in terms of vertical elevation, the great egrets were found between the middle and the top of the trees, whereas the little egrets mainly remained around the middle region. This pattern of vertical stratification is consistent with the observations of other studies, where birds generally align themselves in a vertical direction according to their body size (Hilaluddin et al., 2003; Pang et al., 2019). In addition, the high density of egrets at the middle elevation of the trees is highly attributed to protection from wind/gales, which provides a favourable habitat for inhabiting, breeding and hatching.

The gathered information, including the number of egrets at different periods of time and the spatial distribution of the egrets can be valuable for optimising the management of environmental monitoring and auditing. Construction activities, decommissioning and demolition works and operations can be scheduled at periods with the least bird counts to minimise the associated potential environmental impact on the Penfold Park egretry.

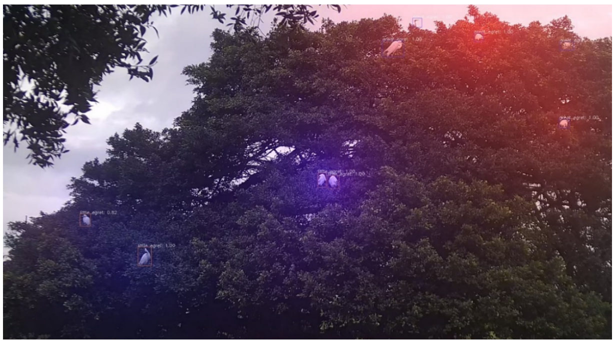


Figure 8. Analysis of the egretry behaviour at the study site from a spatial perspective, where the hotspot intensity represents the density of egret counts at a certain location.

2.6. Limitations and future works

For the current setup with a single Green AI Camera system, the count would be affected when the birds are not within the camera coverage, e.g., birds are fully covered or hidden inside/behind the tree. In the future, these limitations could be tackled via data collection from multiple optimal locations at the monitoring site, whilst considering site factors such as the availability of power supply, installation cost and the impact of the camera system on the overall physical appearance of the site environment.

3. AIBIM Project

3.1. Project background

The DSD is responsible for the management of about 360 public sewerage and stormwater drainage facilities. Building and civil (B&C) assets like structural concrete of the sewerage and drainage facilities, are major assets of the DSD. Very often, these concrete structures are located in aggressive environments and are vulnerable to deterioration due to chemical and biological attacks. Maintenance programmes including regular inspections are thus carried out to ensure that the structures are well maintained in a serviceable state. Manual visual inspection is currently undertaken as the standard practice used in the principal inspections by the maintenance team. However, the practicality of very extensive and frequent inspections is limited by time and labour constraints, as well as occupational safety and health concerns over prolonged work inside sewerage facilities. Therefore, AI technologies have been explored as a feasible tool for enhancing the inspection and maintenance programmes to facilitate the management of B&C assets.

This section on the AIBIM project is organised as follows. The procedures for data annotation of the collected images are described. Details of the two proposed deep learning models for concrete defect identification, using the supervised and unsupervised learning, are presented, focusing on the model architecture, training and evaluation.

3.2. Data annotation

Surveys were conducted at the sewage treatment facilities to collect high-resolution images (4000×6000 pixels) of defects found in concrete structures under a wide range of camera shot angles, lighting conditions and distances from targets. Then, the images were cropped into patches of 256×256 pixels, similar to the input size of deep learning models, to implement data annotation. About 42,000, 6,900 and 6,900 patches were labelled for the class of no defects, cracking and spalling (see Figure 9 for examples). Note that the patches of no defects class represent undamaged concrete surfaces and utilities found inside the sewage treatment facilities, e.g., electrical cables, pipes, and many other items. Other defects, such as leakage and delamination, were rarely observed and not included in the training dataset, i.e., the number of labelled patches of these defects was extremely small.

3.3. Supervised learning for defect classification

ResNet50 (He, 2016) was built and trained as a classifier for concrete defect classification, following the work by Chow et al. (2020a), where details of the model training and evaluation are discussed. The ResNet50 architecture is based on the residual block concept for propagating important features from previous layers to the next layer, in order to overcome the vanishing gradients and to minimise information loss. All cracking and spalling patches, and portions of patches with no defects, i.e., 9,200 out of 42,000 patches, were used to create a balanced dataset for model training. In other words, a total of 23,000 patches, with an approximate ratio of 4:3:3 for the classes of no defects, cracking and spalling, were used. The dataset was then split at a ratio of 4:1 into the training and validation subsets.

Figure 9. Examples of patches for the class of no defects, cracking and spalling.

Similar to the previously described AIEIA project, the model training was implemented using two GeForce RTX 2080 Ti GPUs. Categorical cross-entropy was used as the loss function to facilitate the training of the defect classification. Mini-batch training, with a batch-size of 24, was adopted, and Adaptive Moment Estimation with an initial learning rate of 0.001 was used as the optimisation algorithm (Kingma and Adam, 2015). Several approaches were undertaken to prevent model overfitting, including on-the-fly data augmentation in terms of applying geometric and spectral transformations, batch normalisation (Ioffe and Szegedy, 2015), L2-regularisation and dropout (Srivastava et al., 2014). Model validation was performed at every epoch. The model with the minimum validation loss was considered optimum, after the validation loss stopped decreasing over 20 successive epochs.

Model evaluation was performed to assess the robustness of the trained classifier in relation to various defect attributes. An additional 7,500 patches that had not been seen by the deep learning models were cropped from high-resolution images and annotated, including concrete defects with various degrees of damage found at the walls, beams, slabs and columns with different background conditions, e.g., smooth/coarse concrete surfaces and absence/presence of dirt stains. Figure 10 presents examples of the detected regions of cracking and spalling, highlighted in red and green, respectively. The trained classifier attained an accuracy of 87.1% over three classes, and recall of 86.2% and 89.9% for the class of cracking and spalling, respectively. Further comprehensive evaluation can be found in the documentation compiled by Chow et al. (2020a).

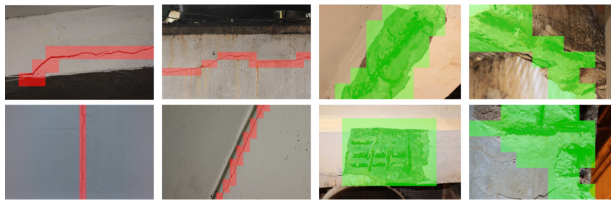


Figure 10. Prediction results generated by the AI-based classifier, where the red and green colours represent the patches of cracking and spalling, respectively (updated from Chow et al., 2020a).

3.4. Unsupervised learning for anomaly detection of concrete defects

As previously described, a large number of patches with no defects, i.e., about 42,000 patches, were annotated. Therefore, these patches with no defects were leveraged to build another deep learning model for concrete defect detection, following the approach implemented by Chow et al. (2020b), where details of the model training and performance can be found. The deep learning model was built as an anomaly detector to screen concrete defects, which were considered as anomalies that deviate from normal instances, i.e., defect-free regions, in terms of visual features.

A convolutional autoencoder was used to achieve this purpose. An autoencoder is capable of learning feature representation of the input and reconstructing the corresponding input without supervision (Géron, 2017), i.e., no labels are required for model training, e.g., the bounding box of egrets in the AIEIA project and the defect labels in the previous AI-based classifier. The bottleneck layer of the convolutional autoencoder, which typically has a much lower dimension than inputs, is used to encode useful representations, rather than merely copy inputs for reconstruction. The convolutional autoencoder is a variant of a typical autoencoder, with the hidden layers comprising multiple convolutional layers designated for the image data, instead of the fully connected layers. In this study, the convolutional autoencoder was trained as a reconstruction-based model using the defect-free patches. This training approach resulted in a poor reconstruction of concrete defects, as a consequence of not learning the relevant anomalous features in advance, and in turn, assisting the detection of concrete defects from images. Such training is also known as unsupervised learning as no label is required. For each pixel of an image, the reconstruction quality was determined using the reconstruction error ( ) computed from the squared difference between the input and the output (reconstructed input) across the three RGB colour channels (Chow et al., 2020b):

$$e = \sum_{c=1}^{3}[p(\mathbf{x},\mathbf{y}) - \hat{p}(\mathbf{x},\mathbf{y})]^2, \quad (2)$$

where p and $\hat{p}$ respectively represent the pixels of the input and reconstructed input, located at row **x** and column **y**; and c is the colour channel of the patch. The input was normalised in the range of -1 to 1, and the hyperbolic tangent was used as the last layer of the convolutional autoencoder to normalise the reconstructed input to the same range. The reconstruction error of all pixels was merged to form an anomaly map delineating the concrete defects.

All the ~42,000 patches with no defects were used for model training. The mean square of the reconstruction error was used as the loss function, and Adaptive Moment Estimation was chosen as the optimisation algorithm to minimise the reconstruction loss. No excessive regularisation was used as the convolutional autoencoder was purposely designated to reconstruct the defect-free patches (i.e., the normal class) as completely as possible.

The trained convolutional autoencoder was tested for both cracking and spalling cases. The patches, cropped from high-resolution images of concrete defects (4000×6000 pixels), were sent to the convolutional autoencoder to generate the reconstructed inputs. Then, anomaly maps of all the patches were generated using the inputs and the reconstructed inputs, and were finally merged to form the entire anomaly map for evaluation.

Table 1. Metric scores attained by the anomaly map and other segmentation methods.

Method	Precision	Recall	F_1 measure	F_2 measure
Cracking case				
Anomaly map (proposed in this study)	0.758	0.879	0.814	0.852
MaxEntropy	0.021	0.879	0.041	0.096
Otsu	0.022	0.766	0.043	0.099
Moments	0.024	0.857	0.046	0.106
Manual	0.463	0.482	0.473	0.478
Spalling case				
Anomaly map (proposed in this study)	0.546	0.843	0.663	0.760
MaxEntropy	0.604	0.484	0.537	0.504
Otsu	0.556	0.390	0.458	0.414
Moments	0.600	0.474	0.530	0.495
Manual	0.551	0.381	0.451	0.406

For comparison, three classical segmentation techniques, i.e., MaxEntropy (Kapur et al., 1985), Otsu (Otsu, 1979) and Moments (Tsai, 1985) methods were chosen to produce segmentation maps of concrete defects. A segmentation map was also manually generated by using a hand-picked threshold, termed "Manual". The original images were transformed to greyscale images prior to segmentation as these methods can only be applied to single-channel images. In addition to visual evaluation, pixel-wise quantitative evaluation was performed with different metrics, including precision, recall, F_1 measure and F_2 measure, where the concrete defects and defect-free regions were considered as the positive and negative class, respectively. Table 1 summarises the metric scores of the anomaly map generated by the convolutional autoencoder and other segmentation methods for the analysed cracking and spalling cases. Note that the scores of precision, recall and F_1 measure of the anomaly map correspond to the highest F_2 measure score.

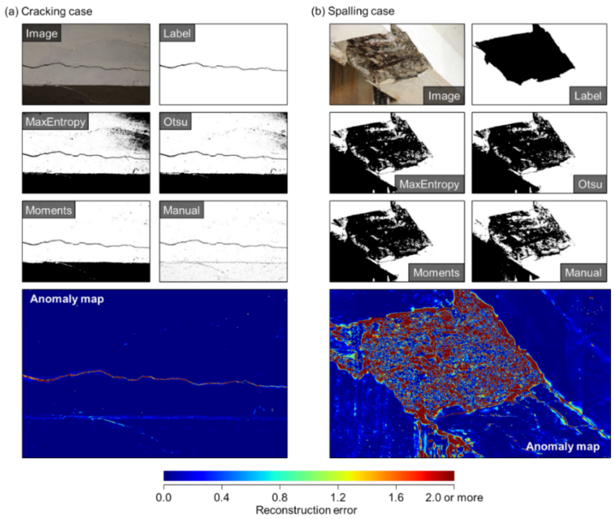


Figure 11. Analysis of the anomaly map generated by the convolutional autoencoder and other segmentation methods: (a) cracking; and (b) spalling case (updated from Chow et al., 2020b).

Figure 11(a) presents the analysis of the cracking case. An elongated crack was found in the middle of the image, and parts of the concrete surfaces are shaded and therefore appear as dark as the crack regions. As a result, over-segmentation occurs, where the shaded region is misidentified as the defect region by the classical methods (false positive). Manual segmentation with a hand-picked threshold is likely to segment the shaded region as the negative class; however, at the same time, under-segmentation occurs with discontinuous segmented cracks. On the other hand, the anomaly map generated by the convolutional autoencoder has relatively high anomaly scores at the crack regions (red shown in the anomaly map of Figure 11(a)), outperforming all other methods in terms of the metric scores (see Table 1). Figure 11(b) presents the analysis of the spalling case, where the outer concrete surfaces were damaged and the reinforced steel bars were exposed. Similar to the cracking case, over-segmentation occurs with most of the shaded regions being misclassified as the foreground by the classical methods (false positive). However, the proposed convolutional autoencoder generates a distinct boundary between the spalled surfaces and defect-free regions (high anomaly scores in red), supported by the high metric scores reported in Table 1. Chow et al. (2020b) can be referred to for a detailed assessment of the proposed anomaly detection for concrete defect detection.

Summarising, the anomaly map generated by the proposed convolutional autoencoder renders a distinct segmentation map, associated with a balanced result without excessive under- and over-segmentation, to facilitate the structural condition assessment of concrete defects. The proposed anomaly map is also appropriate for detecting concrete defects that are rarely observed in the aforementioned data annotation (leakage and delamination), as no knowledge of the relevant defects is required in advance.

3.5. Discussion of future works

AI technology was used to build two deep learning models using supervised and unsupervised learning to detect and classify concrete defects. The well-trained deep learning models for different types of concrete defects can then be integrated with mobile data collection systems and Building Information Modelling (BIM) technologies to formulate a total solution for an AI-ready inspection and management system to implement structural condition assessment of B&C assets. For instance, the detection results could be used to update the existing conditions of structural elements stored in a BIM model with the aid of Light Detection and Ranging (LiDAR) technology. A 3D model of the test site could be built using LiDAR and then calibrated to align with the BIM model. Then, the position of defects could be identified and mapped to respective structural elements in the BIM model, and the associated defect features could be tagged at any empty sections that are not being used by the BIM elements for further review and analysis. Furthermore, the proposed methods can be extended to other structural defects, such as delaminations, water leakage and dislocated structural joints, provided that adequate amounts of images are available.

4. Conclusions

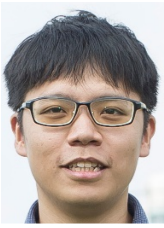
In this paper, two research studies related to the potential application of AI in works projects and routine maintenance by the DSD have been presented, focusing on environmental monitoring (the AIEIA project) and asset management of B&C assets (the AIBIM project). In the AIEIA project, AI technologies were applied to overcome the limitations of manual bird counting, in order to assist the observation of the Penfold Park egretry. The domain randomisation strategy was adopted to augment the data volume for model training and at the same time to force deep learning models to focus on learning the fine-grained bird features used by experts for bird identification. The analysis reveals that the domain randomisation-enhanced model can attain a mean average precision of approximately 90% for bird identification. The AI technologies provide considerable advantages over conventional manual bird counting, where the detection results by the domain randomisation-enhanced model enables bird surveys to be carried out, such as identifying the influx of migratory birds and studying habitat preferences, to aid the formulation of better habitat management and conservation policies. In the AIBIM project, AI technologies were demonstrated to be a feasible alternative to the manual visual inspection of B&C assets. A deep learning model was developed with supervised learning to classify patches taken at the testing site as no defects, cracking and spalling classes, achieving recalls of 86.2% and 89.9% for the classes of cracking and spalling, respectively. At the same time, another deep learning model was built with unsupervised learning, leveraging the abundant amount of patches with no defects, for the anomaly detection of concrete defects. The latter method renders high reconstruction errors that delineate the regions of concrete defects found in images, outperforming the classical segmentation methods, in terms of minimal under- and over-segmentation associated with F_2 measures of 85.2% and 76.0% for the cracking and spalling classes. In short, the advanced AI-enabled analytics enable efficient transformation of raw data into structured insights for proactive decision-making in environmental monitoring and asset management, thereby strengthening the DSD's delivery of its vision of providing world-class wastewater and stormwater drainage services.

Acknowledgements

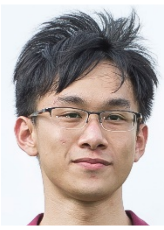
The two research studies were supported by the Drainage Services Department of the Hong Kong Special Administrative Region Government, the Research Talent Hub for the Hong Kong Innovation and Technology Fund, the Hong Kong Research Grants Council (project no. 16205021) and The Hong Kong University of Science and Technology Sustainable Smart Campus as a Living Lab project. The authors wish to express their gratitude to the Drainage Services Department of the Hong Kong Special Administrative Region Government for permission to publish this paper. However, the views expressed are those of the authors and do not necessarily reflect government policy in any way.

Notes on contributors

Dr Jun Kang Chow currently is a Postdoctoral fellow at the Department of Civil and Environmental engineering, HKUST and is also a founding member of the DESR Lab. He received a scholarship for his undergraduate study at the HKUST in 2011, majoring in Environmental Engineering, and in 2014 was a semi-finalist in the HKUST President Cup, the highest honour for undergraduate research. He was awarded his M.Phil. degree in Civil and Environmental Engineering in 2016 and also won the Best Master Thesis Award 2018, ASCE Hong Kong Section. He received the Hong Kong Ph.D. fellowship for his Ph.D. study, obtained his degree in 2019 and received the Ringo Yu best Ph.D. award in 2020. His research interest is applying different deep learning techniques to tackle the current challenges in civil and environmental engineering and he has published many relevant papers, from clay fabric detection at the microscale, bird identifications and classifications for environmental impact assessment, and the automation of landslide inventory, to concrete defect detection and classification.

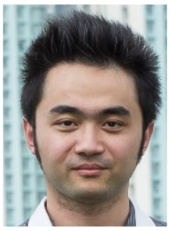


Mr Pin Siang Tan is a founding member of the Data-Enabled Scalable Research (DESR) Lab. He is currently a Ph.D. student under Prof Yu-Hsing Wang. He was awarded the Master of Philosophy degree in Civil and Environmental Engineering in 2015 for the research "A scalable architecture for continuous, in-time landslide early warning and monitoring" which precipitated the eventual technologies adopted in the DESR Lab. He also won the Third Class Award in Freescale Smart Car competition and is proficient in embedded hardware and software development and mixed-signal PCB layout. He is the core member of several research projects, covering the development of the Smart Soil Particle (SSP) sensor node module and a scalable Big Data server farm for high throughput data ingestion and distributed processing to handle the massive amount of data from SSPs, as well as implementation of AI-enabled data analytics in different civil and environmental engineering applications.



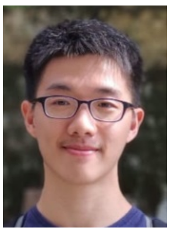
Mr Kuan-fu Liu is a member of the DESR Lab, HKUST. He was awarded the Master of Philosophy degree in Technology Leadership and Entrepreneurship for the research “Mobile Vision System for Automatic Building Defects Detection” under the supervision of Prof Wang in 2020. He is the team member responsible for developing the Green AI Camera in the AIEIA project, covering installation and maintenance of all the hardware and software components for data acquisition, transfer and storage.

Dr Xin Mao is a member of the DESR Lab, HKUST. He was awarded the Doctor of Philosophy degree in Civil Engineering for the research “Expert Knowledge Guided Deep Learning for Civil and Environmental Engineering Applications” under the supervision of Prof Wang in 2021. He is the team member responsible for developing computational algorithms for bird detection and the associated analysis of egretry behaviour in the AIEIA project. He studied and applied the domain randomisation strategy, via developing virtual egret models, to resolve the limited amount of labelled bird data, so as to enhance the capability of deep learning models in learning fine-grained bird features used by experts for bird identification.



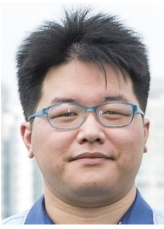
Mr Zhaoyu Su is a member of the DESR Lab, HKUST. He is currently a candidate of Doctor of Philosophy degree in Civil Engineering for the research “Advanced 2D and 3D Computer Vision for a Smart City – from Image to Point Cloud” under the supervision of Prof Wang. His research interest is developing computational algorithms for advancing the model accuracy and efficiency of deep learning models, particularly convolutional neural networks, in different civil and environmental engineering applications, such as landslide segmentation and vehicle detection.

Mr Ghee Leng Ooi is a member of the DESR Lab, HKUST. He won the HKUST President's Cup for Research Excellence in 2009 and was awarded the Master of Philosophy degree in Civil and Environmental Engineering for the research “Exploratory study on initiation mechanism of flow landslide in loose soil” under the supervision of Prof Wang in 2012. He completed an exploratory study on the landslide initiation mechanism in loose slopes under Prof Wang in 2012. He recruited the current team members and works hand in hand with Prof Wang and Mr Tan Pin Siang to bring big data analytics to fruition. In 2012 and 2013, he led the team members into the final rounds of the HKUST Entrepreneurship Competition by presenting the business plans on Smart Soil Particles.



Dr Ye Hur Cheong is a member of the DESR Lab, HKUST. He was awarded the Doctor of Philosophy degree in Industrial Engineering and Logistics Management for the research “Vertical Stripes within Chinese Characters Slow Down Reading Speeds: A Binocular Fixation Issue”. He is the team member responsible for the analysis of Penfold Park egretry behaviour in the AIEIA project, focusing on the bird count of different species, behavioural study, spatial preferences and inter- and intra-specific interaction.

Ms Mei Ling Leung is a member of the DESR Lab, HKUST. She was awarded the Bachelor of Engineering degree in Electronic and Computer Engineering. She is responsible for designing all the hardware components of the AIEIA and AIBIM projects, including fabricating the printed circuit boards, assembling electronic components and installing the Green AI Camera at the testing site.



Mr Jimmy Wu is a member of the DESR Lab, HKUST. He is currently a candidate of the Doctor of Philosophy degree in Civil Engineering for the research "Unsupervised Anomaly Detection and Clustering of Multi-dimensional Time Series Data for Smart Cities Application" under the supervision of Prof Wang. He is responsible for analysing the performance of the deep learning models for concrete defect detection in the AIBIM project.

Ir Hok Man Chan is currently a senior engineer with the Drainage Services Department. He graduated with a Bachelor's degree in Civil Engineering and a Master of Science degree, both from The University of Hong Kong. He has about 20 years of experience in a variety of projects, including highways, railways, transport planning, stormwater drainage and sewerage. He has been serving in the Drainage Services Department since 2013 and has been involved in various drainage improvement works, river revitalisation, sewage treatment and landscaping works. He is currently responsible for implementing the relocation of Sha Tin Sewage Treatment Works to caverns project.

Ir Lai Yuk Yip is currently a senior engineer with the Drainage Services Department. She graduated with a Bachelor's degree and a Master of Philosophy in Civil and Structural Engineering from The Hong Kong University of Science and Technology. She has about 18 years of experience in a variety of projects, including geotechnical, highways, stormwater drainage and sewerage works. She has been serving in the Drainage Services Department since 2012 and has been involved in various drainage improvement works, sewerage treatment and cavern construction works. She managed the AIEIA project in applying the Green AI system in the project of Relocation of Sha Tin Sewage Treatment Works to caverns, with the aim of identifying and analysing egrets near the project site for the planning of effective measures to minimise the impacts to egrets during the construction period.

Ir Ka Chun Chow has over 15 years of relevant experience and 10 years of post-qualification experience in civil engineering. He has been involved in the investigation, design, tender and construction stage for numerous infrastructure projects involving vehicular underpasses, bridges, subways, retaining walls, noise barriers, pumping stations, service reservoirs, road and drainage, sewerage, mainlaying and landscape works. He joined the Government of HKSAR as an engineer in 2013. When he was working in the Civil Engineering and Development Department, he was mainly devoted to the management of infrastructure projects at the north apron of the Kai Tak Development. He was then posted to the Drainage Services Department (DSD) in 2017 and was assigned to develop the strategy and applications of the Building Information Modelling (BIM) projects. He was a key member developing the BIM management system for the DSD and fostering the collaboration with HKUST to explore the potential of integrating state-of-the-art technologies such as artificial intelligence into BIM with a view to improving the overall operation and maintenance efficiency in routine structural inspections.

Prof Yu-Hsing Wang received his B.S. and M.S. degrees in Civil Engineering from National Taiwan University and Ph.D. in Civil Engineering from Georgia Institute of Technology where he received the George F. Sowers Distinguished Graduate Student Award for Ph.D. Students. Prof Wang has been a Professional Geotechnical Engineer in Taiwan since 1996. Currently, he is Professor in the Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology (HKUST). In 2005, he received the ASTM International Hogentogler Award. In 2008 and 2017, he received the School of Engineering Teaching Award, HKUST. In 2013, he received the Distinguished Alumni Award from the Department of Civil Engineering, National Taiwan University. He has been invited to Keynote, Plenary and Theme lectures in international conferences, including IEEE conferences. Prof Wang is also the founder and director of the Data-Enabled Scalable Research Laboratory (DESR Lab), HKUST. The mission of the DESR Lab is to innovate smart city infrastructure to connect citizens and decision-making, specialising in the applications of Vertical AI, integrated with Geotechnical Internet of Things (Geo-IoT), Big Data Analytics, and Deep Learning, etc., for sustainable urban development and city resilience. The DESR Lab has been working closely with industry and government departments, such as CEDD, GEO, EPD, Transport Department, HKO, DSD, WSD, Lands Department, Development Bureau, AFCD, etc. The research developments from the DESR Lab have paved a smooth way for the Hong Kong government to spearhead smart city development.

References

[1] Bengio Y (2009). Learning Deep Architectures for AI. *Foundations and Trends®, Machine Learning*, 2(1), pp. 1–127.

[2] Carey G (2001). *The Avifauna of Hong Kong*. Hong Kong: Hong Kong Bird Watching Society.

[3] Choi Y-S, Kwon I-K and Yoo J-C (2007). Foraging Habitat Preferences of Herons and Egrets. *Journal of Ecology and Environment*, 30, pp. 237–244.

[4] Chow JK, Su Z, Wu J, Li Z, Tan PS, Liu KF, Mao X and Wang YH. (2020a). Artificial intelligence-empowered pipeline for image-based inspection of concrete structures. *Automation in Construction*, 120, pp. 103372.

[5] Chow JK, Su Z, Wu J, Tan PS, Mao X and Wang YH (2020b). Anomaly detection of defects on concrete structures with the convolutional autoencoder. *Advanced Engineering Informatics*, 45, pp. 101105.

[6] He K, Zhang X, Ren S and Sun J (2016). *Deep residual learning for image recognition*. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, pp. 770–778.

[7] Hilaluddin J, Shah N and Shawl T (2003). Nest Site Selection and Breeding Success by Cattle Egret and Little Egret in Amroha, Uttar Pradesh, India. *Waterbirds*, 26, pp. 444–448.

[8] Ioffe S and Szegedy C (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. In: Proceedings of the 32nd International Conference on Machine Learning. Lille: MLResearch Press, pp. 448–456.

[9] Kapur JN, Sahoo PK and Wong ACK (1985). A new method for gray-Level picture thresholding using the entropy of the histogram. *Graphical Models and Image Processing*, 29(3), pp. 273–285.

[10] Kingma DP and Ba J (2015). *Adam: A method for stochastic optimization*. In: 3rd International Conference for Learning Representations. [online]. San Diego: DBLP Computer Science Bibliography. Available at https://arxiv.org/abs/1412.6980. [Accessed on 12 July 20201].

[11] Lin T-Y, Dollar P, Girshick R, He K, Hariharan B and Belongie S (2017). *Feature Pyramid Networks for Object Detection*. In: Proceedings of the 2017 IEEE conference on computer vision and pattern recognition. Honolulu: IEEE, pp. 2117–2125.

[12] Mao X, Chow JK, Tan PS, Liu KF, Wu J, Su Z, Cheong YH, Ooi GL, Pang CC and Wang YH (2021). Domain randomization-enhanced deep learning models for bird detection. *Scientific Reports*, 11(1), pp. 1–3.

[13] Otsu N (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), pp. 62–66.

[14] Pang RH, Yu TL and Busam M (2019). Low Breeding Success of The Little Egret (Egretta garzetta) near Residential Areas and in Colonies Exposed to Gales: A Comparison of Colony in Sichuan, Southwest China, with Literature. *Animal Cells and Systems*, 23(3), pp. 235–240.

[15] Peng XB, Andrychowicz M, Zaremba W and Abbeel P (2018). *Sim-to-Real Transfer of Robotic Control with Dynamics Randomization*. In: 2018 IEEE International Conference on Robotics and Automation. Brisbane: IEEE, pp. 3803–3810.

[16] Ren S, He K, Girshick R and Sun J (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), pp. 1137–1149.

[17] Sadeghi F and Levine S (2017). *CAD2RL: Real Single-image Flight without A Single Real Image*. In: Robotics: Science and Systems 2017. [online]. Cambridge: MIT LIDS. Available at: https://rss2017.lids.mit.edu/static/papers/13.pdf. [Accessed on 12 July 2021].

[18] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), pp. 1929–1958.

[19] The Hong Kong Bird Watching Society (2018). *The First GPS-Tracking Research of Egrets in Hong Kong Fishpond Plays A Key Role for Both Resident and Migratory Little Egret*. [online]. Available at https://cms.hkbws.org.hk/cms/en/hkbws/egret-tracking-2. [Accessed on 12 July 2021].

[20] Tobin J, Fong R, Ray A, Schneider J, Zaremba W and Abbeel P (2017). *Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World*. In: Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vancouver: IEEE, pp. 23–30.

[21] Tsai W (1985). Moment-preserving thresholding: a new approach. *Computer Vision, Graphics, and Image Processing*, 29, pp. 377–393.

[22] Young L (1998). The Importance to Ardeids of The Deep Bay Fish Ponds, Hong Kong. *Biological Conservation*, 84(3), pp. 293–300.