A New Adopted YOLOv9 Model for Detecting Mould Regions Inside of Buildings

Taha Mansouri[1], Md. Shadab Mashuk[1], Shivakumara Palaiahnakote[1], Aaron Chacko[2], Lawrence Sykes[2], and Ali Alameer[1]

[1]School of Science, Engineering and Environment, University of Salford, Salford, Greater Manchester, M5 4QJ, UK. Email: t.mansouri@salford.ac.uk, m.s.mashuk@salford.ac.uk, s.palaiahnakote@salford.ac.uk, a.alameer1@salford.ac.uk.

[2]Copious Group Limited, Rochdale, UK. Email: dev@copious.co.uk, info@copious.co.uk

Abstract

Molds on wall and ceiling surfaces in damp indoor environments especially in houses with poor insulation and ventilation are common in the UK. Since it releases toxic chemicals as it grows, it is a serious health hazard for occupants who live in such houses. For example, eye irritation, sneezing, nose bleeds, respiratory infections and skin irritations. Furthermore, there are chances of developing serious medical conditions like lung infections and respiratory diseases which may even lead to death. The main challenge here is that due to their irregular patterns, camouflaged with the background, it is not so easy to detect with our naked eyes in the early stage and often confused as stains. Therefore, inspired by the accomplishments of the Yolo architecture for object detection, the Yolov9 model is explored for mould detection by considering mould region as an object in this work. The overall result shows a promising 76% average classification rate. Since the mould does not have a shape, specific pattern or colour, adapting the Yolov9 for accurate mould detection is challenging. To the best of our knowledge, this is the first of its kind compared to existing methods. Since it is the first work, we constructed a dataset to perform experiments and evaluate the proposed method. To demonstrate the proposed method's effectiveness, the results were also compared with the results of the Yolov8 and Yolov10 models.

Keywords: Mould, Computer Vision, Deep Learning, Indoor Health, Mould detection, Mould classification.

1. Introduction

The presence of damp and mould growth in buildings can cause serious health problems for residents who are exposed to them. These problems include respiratory infections, allergies, eczema, and bronchitis (Menneer et al., 2022; Sharpe et al., 2016; Simon-Nobbe et al., 2008). The severity of mould exposure can cause pronounced side effects in vulnerable populations, particularly amongst children, the elderly, and those in care homes (Du et al., 2021) with occurrences of deaths linked to mould exposure has been reported in the UK recently (Caillaud et al., 2018; Jones & O'Donoghue, 2023; Jubayer et al., 2021). Beyond health concerns, mould can also damage building infrastructures; walls, paints and frames and also reduce their energy efficiency (Song et al., 2017). Professional assessment by experienced surveyors remains the benchmark for detecting mould and its impact on buildings. However, it can be time-consuming and expensive, and some studies suggest it can be subjective and miss early signs of mould growth (Menneer et al., 2022). Swab sampling and laboratory analysis, another traditional method, offer more accurate identification of different variations of mould types but also suffer from delays and high costs (Hamilton et al., 2015).

Mould and dampness pose significant concerns in the daily lives of people and in the UK, alone approximately 31% of homes are affected by dampness issues due to poor ventilation and insulation, elevating the likelihood of condensation (Menneer et al., 2022). The existence of indoor dampness, at-attributed to water ingress, rising dampness, and condensation, can result in heightened mould contamination. Even housing interventions aimed at mitigating fuel poverty through energy efficiency measures may be susceptible to condensation and mould contamination unless proper ventilation and heating are in place (Hamilton et al., 2015; Sharpe et al., 2016). All of these factors influences are important to consider because the presence of mould or mouldy odour has an adverse effect on health for a wide range of conditions (Simon-Nobbe et al., 2008), including the development and/or exacerbation of a range of allergic and non-allergenic diseases (Menneer et al., 2022).

In the same way, due to eyes and skin irritations, people lose focus and concentration, which leads to poor performance and productivity at the workplace as well as at home (Simon-Nobbe et al., 2008). Therefore, one can argue that overall, it affects the growth of the economy. Hence, there is an immense need to develop a model to detect mould in the early stages to prevent the loss of life and other major health issues. Sample images containing moulds of different patterns on walls can be seen in Fig. 1, with the third image, showing how difficult it can be to locate mould with the naked eye while for the first and second, we can locate but the background complexity varies. The samples also confirm that the mould does not have a shape, pattern or specific colour which is embedded in the background with varying complexities.

Fig. 1. Sample Images of Diverse Mould Patterns on Various Surfaces

There are solutions in the market for removing the mould on the wall and other ceiling surfaces, but these solutions can be used after detecting and locating the mould. It is also true that these solutions do not stop the recurrent growth of mould permanently and therefore, we see them coming back again and again within a short period. This situation demands a predictive model that can automatically detect them from images when mould develops on walls and other places in the house. This shows that a solution alone is not sufficient to solve the problem of mould. So, the detection of mould automatically is necessary to develop automatic systems like robots. The robot can claim the wall to detect and remove the mould in the house and workplace without human intervention and interference (Li et al. 2023). This is a major application and significance of the proposed work, which has great practical implications and market in the business.

When we look at the literature and state-of-the-art, there are no methods for mould detection to the best of our knowledge. However, one can find the technology to identify and detect specific features, defects, and objects, including potential signs of mold growth in buildings (Damp and Mould: Understanding and Addressing the Health Risks for Rented Housing Providers, 2023). Motivated by the work (Damp and Mould: Understanding and Addressing the Health Risks for Rented Housing Providers, 2023), where fungus on food was detected by the computer vision method (Nair et al., 2021), we propose a computer vision model for mould detection in this work. As such we believe this is the first work which aims to automatically detect moulds of any pattern at different stages of its growth. Therefore, inspired by the success of Yolo architecture for object detection in adverse situations including tiny objects in complex backgrounds (Wang et al., 2024; Wang et al., 2024), we have explored the YOLOv9 for mould detection in this work. The challenge in this study has been to look at how to make the object detection model work on mould detection using adopted YOLOv9. In addition, we create a new dataset with clear annotations for experimentation, evaluation in this study and released to the public to support research reproducibility to investigate further on mould detection, prediction and identification. Therefore, dataset creation and annotation is considered as one more key contribution of the proposed work.

The structure of the paper is organized as follows. The method related to fungus detection and other related to mould detection are reviewed in Section 2. In Section 3, the overview of YOLO and the details of YOLOv9 are presented. Experimental analysis and the results are discussed in Section 4. The findings are listed in Section 5.

2. Related Work

Since there are no existing methods for mould detection, we review the related methods of defect, cracks pothole detection in this section. This is because the idea of defects, cracks and potholes are the closest available to mould detection. In particular, the methods usually check sudden changes in the background for suspecting unusual regions in the images.

Traditional inspection methods involve the engagement of building surveyors to conduct a comprehensive condition assessment (Kong et al., 2018; Noel et al., 2017; Song et al., 2017). This assessment typically entails a time-consuming on-site inspection during which they systematically document the physical state of building components through the use of photographs, note-taking, drawings, and information provided by the client (Annamdas et al., 2017; Kong et al., 2018; Noel et al., 2017; Song et al., 2017).

This assessment typically entails a time-consuming on-site inspection during which they systematically document the physical state of building components through the use of photographs, note-taking, drawings, and information provided by the client (Annamdas et al., 2017; Kong et al., 2018; Noel et al., 2017; Song et al., 2017). The collected data are subsequently subjected to thorough analysis to generate a comprehensive report summarizing both the condition of the building and the state of its elements (Lorenzoni et al., 2016). Furthermore, this information serves as the foundation for estimating immediate and projected long-term costs associated with building renewal, repair, and maintenance (Oh et al., 2017). These estimates empower facility managers to address their current operational needs while enhancing their ability to forecast real estate portfolio renewal and allocate funding for capital projects (Oh et al., 2017).

In summary, it's worth noting that the current procedures for assessing asset conditions are not only highly time- consuming and labor-intensive but also come with safety concerns, particularly when surveyors are required to access challenging locations such as rooftops (Mimura & Mita, 2017; Mita, 2016). Therefore, this demands an automatic system for mould detection and removal without human intervention.

In recent years, researchers have explored the application of soft computing and machine learning-based techniques to automate asset condition inspection. Noteworthy efforts encompass structural health monitoring using Bayesian methods (Zhang et al., 2019), surface crack estimation using Gaussian regression, support vector machines (SVM), and neural networks (Davoudi et al., 2018), SVM for recognizing wall defects (Hoang, 2018), crack detection on concrete surfaces with deep belief networks (Pahlberg et al., 2018), detection of cracks in oak flooring using ensemble methods like random forests (Valero et al., 2019), and deterioration assessment using fuzzy logic (Pragalath et al., 2018). Furthermore, research in defect detection has extended to infrastructural assets such as road surfaces, bridges, dams, and sewerage pipelines. As seen in (Valero et al., 2018) use of computer vision and machine learning models to detect defects in historical buildings with masonry walls. The use of neural networks to identify cracks and crack types from pavement images can be seen in (Lee & Lee, 2004). The authors used three different architecture types, compared and validated using actual pavement crack images.

Similar work can be seen in (Koch & Brilakis, 2011) to identify potholes in pavements and potential defects using the automated approach of automated analysis and computer vision techniques. A generalized machine learning approach using the Adaboost algorithm to detect defective road surfaces from images can be seen in (Jahanshahi et al., 2013). They use linear and nonlinear filters to identify textural patterns using a generic supervised learning model. Similar work can be seen in (Abdel-Qader et al., 2003) again using a supervised classifier known as Semantic Texton Forests (STFs) algorithm and identified two types of cracks, longitudinal and transverse, patches and potholes. The use of a depth sensor applying the RGB-D method can be found in (Radopoulou & Brilakis, 2017) to identify road conditions and defects in a cost-effective manner.

Furthermore, the methods (Giakoumoglou et al., 2023; Perez et al., 2019; Soeb et al., 2023) used YOLO for detecting sooty mould and other diseases from plant leaves. The use of learning using multi-spectral images can be seen in (Perez et al., 2019) and another one investigated anomaly detection in cosmetics labelling during the manufacturing process (Soeb et al., 2023). However, the objective and scope of the methods are different from mould region detection as part of the proposed work.

Overall, when we look at reviews on different methods on different topics related to mould detection, none of the methods aims to detect mould. Although mould detection has practical implications and impact, the above topics seldom address this issue. This indicates that detecting mould successfully is challenging and complex. In addition, different object detection models have been proposed for identifying fungus but not mould. Therefore, this work for the first time adopts YOLOv9 for detecting moulds of any pattern with any background using a large real-world dataset of mould collected from households in the UK affected by this problem.

3. Proposed Methodology

In this paper, the focus is to detect the mould while differentiating from other stains present in the images, which may typically be confused with the mould when we look at it with the naked eye. Therefore, detecting mould accurately is complex and challenging. It is known that the YOLO can detect objects in complex backgrounds and other adverse situations in real-time (Wang et al., 2024; C.-Y. Wang et al., 2024). This has motivated us to explore the YOLOv9 for mould detection in this work. However, there are many YOLO variants which belong to the YOLO family including YOLOv8 and the latest; YOLOv10. Out of all YOLO including YOLOv10, we prefer to choose YOLOv9 for mould detection in this work. The reason is that YOLOv9 model training and validation is simple and effective with a small number of samples compared to YOLOv8 and YOLOv10. Since there is no standardised dataset for mould detection, we create our dataset along with the ground truth. Therefore, one cannot expect many samples for training and testing as it involves high labor costs and time. It is noted that YOLOv10 requires more samples for training because of its advanced architecture compared to YOLOv9. If YOLOv10 does not train on enough samples, it does not work well for any problem. Similarly, YOLOv8 also improves over its predecessors in terms of accuracy, speed and deployment offering flexibility across CPUs, GPUs and other edge devices, it has serious limitations when it comes to highly specialized and unique datasets and relies heavily on the quality and quantity of the data. Therefore, the proposed work prefers YOLOv9 as it does not require a large number of samples for training. In addition, YOLOv9 extracts effective features with a small number of samples. These observations motivated us to choose YOLOv9 for adoption to obtain results for mould detection although we still compare and evaluate the performance with a couple of other YOLO models; YOLOv8 and YOLOv10 to be specific.

The block diagram to understand the process of adaptation of YOLOv9 can be seen in Fig. 2. The Roboflow is used to fix the bounding box to represent the mould region. It calculates the centroid and corners of the bounding boxes using the X and Y coordinates of pixels of the mould region to fix the rectangular bounding box.

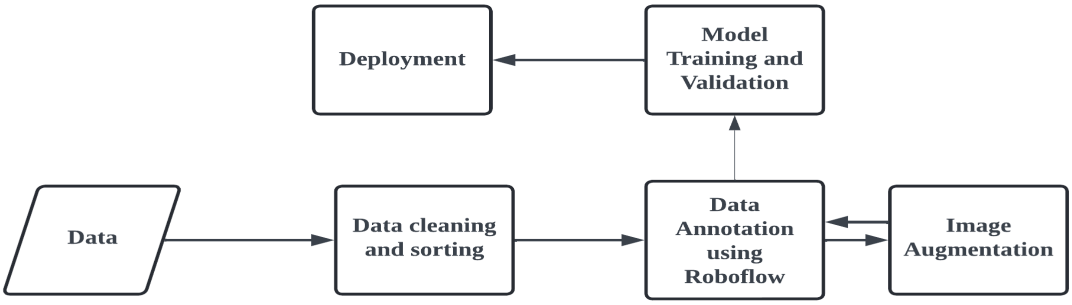


Fig. 2. The logical flow of the proposed method for mould region detection

3.1. Overview of YOLO - YOU ONLY LOOK ONCE

Object detection typically involves locating objects of interest within images or video with major applications ranging across different sectors such as surveillance, robotics, and self-driving cars to name a few. There are two types of approach - single-shot object detection and Two-shot object detection. Single shot detection, as the name suggests, typically processes the classification and detection in one pass making it fast and efficient for real-time object detection. The downside is it can be less accurate, particularly for smaller objects. Two-shot detection processes the image in two passes. The first pass makes a potential decision on object locations within the image and the second pass refines the decision before reaching a final prediction.

YOLO is a single-shot object detection algorithm using an end-to-end CNN model and following the architecture in Fig. 3. The image is divided up into S×S grids and each of the grids will be defined according to the vector in Equation (1). Here (P_c) is the probability of the object class, B_x and B_y are coordinates of the center of that bounding box, relative to the cell, B_x and B_y are the width and height of the bounding box relative to the whole image. Each cell predicts B bounding boxes and probability scores for those boxes during training. These scores reflect the confidence of the model that the grid contains an object and how accurate it thinks the predicted object is within the bounded box. It uses Equation (2) to calculate the IoU between the prediction and the bounding boxes which is the ground truth during training by assigning a single predictor for each object. It also uses a technique called non-maximum. suppression (NMS) to avoid confusion while identifying IoU for the same object but having overlap at different positions. In this study, we explore YoloV9 for mould detection irrespective of background complexities, colour and in shape.

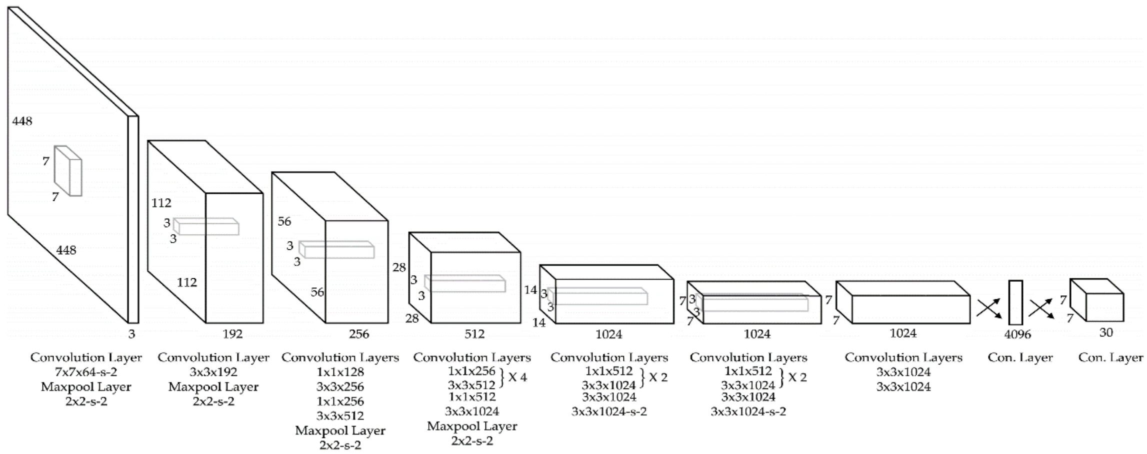


Fig. 3. Original architecture of the Yolo model

3.2. Adopted YOLOv9 for Mould Region Detection

This latest version of the YOLO object detection model, YOLOv9, offers several advantages over its predecessor, including programmable gradient information (PGI) and the generalized efficient layer aggregation network (GELAN) (Wang et al., 2024) . It supports both CPU and GPU devices through a new API, enabling faster training. By integrating PGI and the GELAN architecture, as shown in Fig. 4 and Fig. 5, YOLOv9 significantly enhances the model's learning capabilities and ensures robust retention of information when transferring data across the layers of the feedforward architecture.

This improvement is rooted in the principles of the information bottleneck and reversible functions. During training, as data passes through successive layers of deep neural networks, information loss is common. The PGI feature counters this by ensuring reliable gradient generation, leading to improved model convergence and overall performance. Additionally, the concept of reversible functions allows for more accurate updating of model parameters during training, reducing information degradation and resulting in more accurate object detection.

The effectiveness of YOLOv9 for mould detection can be seen in Fig. 6, where for images of different mould patterns, the feature maps obtained by YOLOv9 indicate the presence of mould in different backgrounds. It is also observed from Fig. 6 that the proposed adopted YOLOv9 can work well for visible mould in the first image and invisible mould in the second and third images. This is the main advantage of the proposed adopted YOLOv9 compared to other models developed in the past.

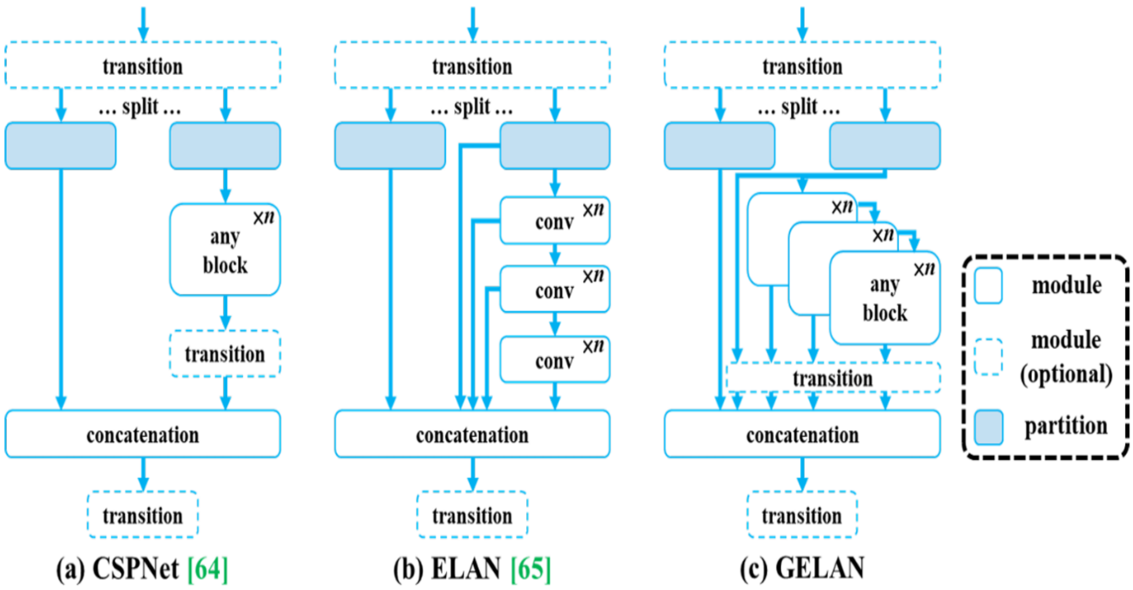


Fig. 4. Architecture of the GELAN component of YOLO v9 (Wang et al., 2024)

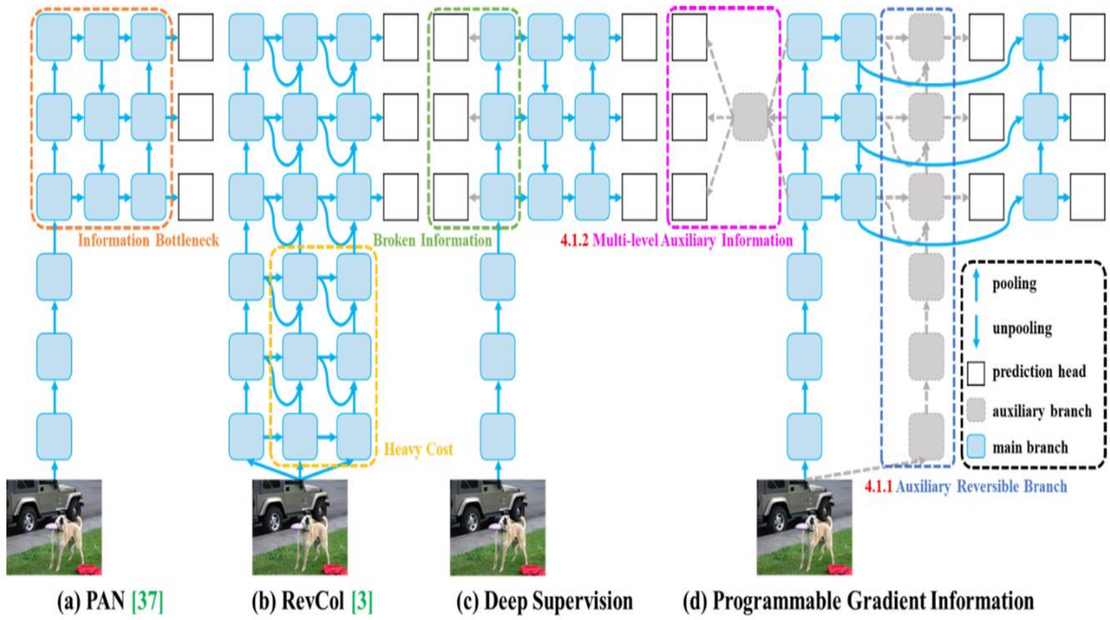
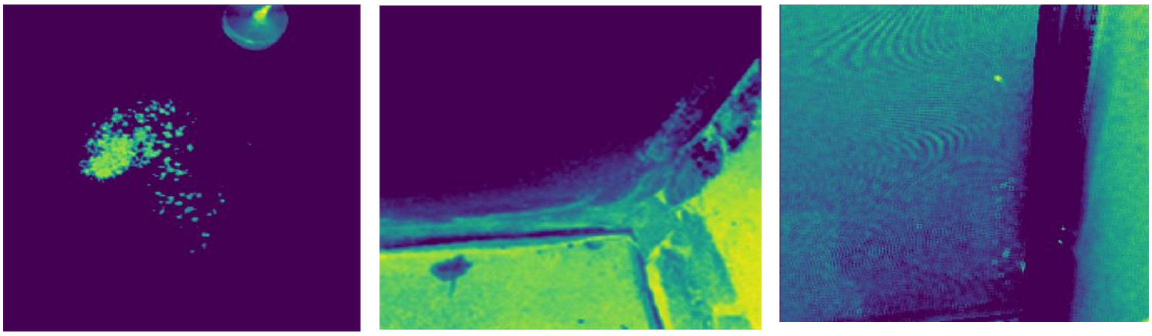


Fig. 5. Architecture of the PGI component of YOLO v9 (Wang et al., 2024)

(a) Sample images containing different mould patterns on the wall



(b) Features maps for the images shown in (a)

Fig. 6. The mould region is highlighted in the feature space.

4. Experimental Results

As mentioned earlier, we collected our own dataset for experimentation and evaluation of the proposed method using images provided by our collaborators at Copious Ltd. For training and testing, we manually annotated the mould of different patterns and considered bounding boxes as ground truth for training and performance evaluation. Since there are no established mould detection methods in the literature, we compared the proposed YOLOv9 model with the baseline YOLOv10 to demonstrate its superiority.

4.1. Data Annotation and Evaluation

Copious Ltd, the partner company of this research, provided a dataset of 5,324 images collected from client disrepair claims, showcasing various types of damage. From this dataset, we extracted images related to mould as the target class, resulting in a total of 2,359 images for our study. As the goal was mould detection, a few samples containing other damage types are shown in Fig. 7, where one can see mould with blurry visuals, poor lighting, and tilted angles. Therefore, our dataset includes sample images of all possible variations in mould formation which are common in the house, especially in the UK. The dataset and annotations can be considered as a benchmark for further study on mould detection, prediction and identification. Sample annotations for different mould regions can be seen in Fig. 8, where we fix manually bounding boxes. This task is not trivial because mould does not have the shape to draw a bounding box. To alleviate this problem, we locate dense mould pixels as centres to estimate the approximate boundary to cover most of the mould pixels. However, the rectangle is not the right choice for labelling annotations because of arbitrary patterns and scattered mould pixels in multiple directions. Therefore, there is a scope for better annotations possibly annotating pixel by pixel to define foreground and background. This is beyond the scope of the proposed work and will be considered as our future work. We plan to release the dataset to the public to support research production if the company permits.

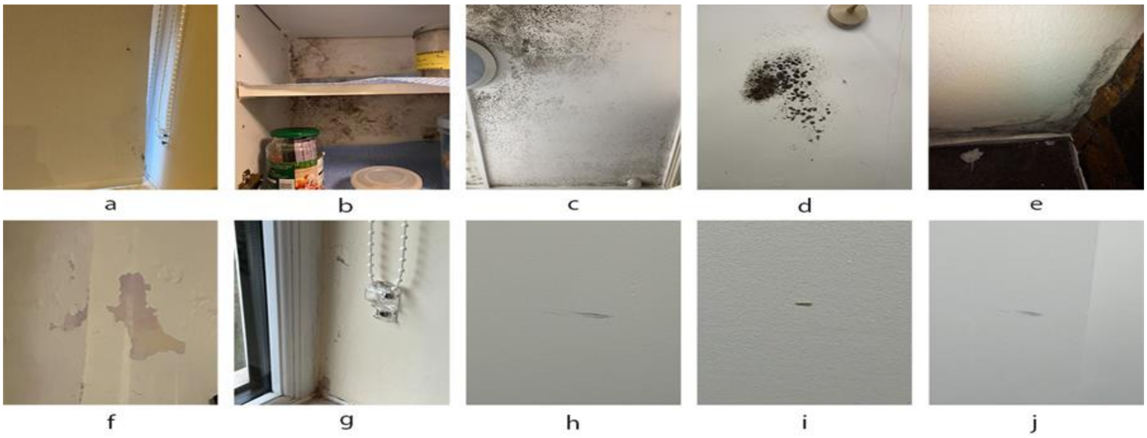


Fig. 7. Sample Images from the Mould Detection Dataset. a-e: Examples of positive class images containing patches of mould on various surfaces. These images are taken under different lighting conditions and at varying angles. f-g: Examples of negative class images depicting color deterioration and paint peeling. These are not classified as mould by the model. h-j: Examples of negative class images displaying stains on surfaces. These are also not classified as mould by the model.

Fig. 8. Annotated and Preprocessed Mould Detection Dataset. This figure displays the result of preprocessing the dataset for training the mould detection model. Bounding boxes (rectangles) are drawn around individual patches of mould (object detection task) within the images. Since some images contain multiple areas of mould, several bounding boxes might be present. It's important to note that the images might have undergone preprocessing steps before this annotation stage.

Building upon the dataset, to measure the performance of the proposed model, we consider well-known standard metrics, namely, *precision*, *recall*, *F1-score* and confusion matrix. The measures provide insights into the models' ability to correctly identify objects (true positives) and minimize false positives and negatives. Additionally, we leverage Mean Average Precision *mAP* to offer a single, consolidated measure of the models' overall detection accuracy across all object classes present in the dataset. Finally, Intersection over Union *IoU* is employed to evaluate the accuracy of bounding box predictions, quantifying the degree of overlap between the predicted and ground truth bounding boxes. For the experiments, the dataset is split up into training, validation and test sets; 2063, 196 and 100, respectively.

YOLOv8 (Sohan et al. 2023): The model released before YOLOv9 improves significantly over its previous versions in terms of speed, accuracy and precision and supports multiple hardware platforms such as CPU, GPU and edge devices. It proposes a unified framework for object identification, localization and classification to detect in complex scenarios even with occluded objects. Unfortunately, because of this support, it is computationally resource-intensive and has limitations when it comes to nonstandard or smaller datasets. The base model may also find the detection of smaller objects from images when they are cluttered or partially occluded, thus impacting overall accuracy.

YOLOv10: This model was released after we conducted our study, and we decided to compare our results with it. Developed by researchers at Tsinghua University, it addresses several issues related to post-processing and model architecture that previously impacted end-to-end object detection, introduced latency, and affected overall accuracy. The non-maximum suppression (NMS) technique which is a trademark of Yolo model has a significant computational overhead which has now been removed and replaced with Consistent Dual Assignments for NMS-free Training strategy. Previous work with NMS during training suppressed redundant predictions due to multiple overlapping IoU resulting in suboptimal performance. The new Yolov10 proposed dual label assignments by combining one-to-one and one-to-many labels during training. The proposed changes for an optimal post-processing technique can be seen in Fig. 9. Furthermore, the YOLOv10 has significant changes in the overall architecture by introducing "Holistic Efficiency-Accuracy Driven Model Design" design changes such as Lightweight classification head, Spatial-channel decoupled downsampling, Rank-guided block design, Large-kernel convolution and Partial self-attention (PSA) (Wang et al., 2024). Therefore, it is heavy compared to YOLOv9 and YOLOv10 requires more samples to achieve the best performance compared to YOLOv9.

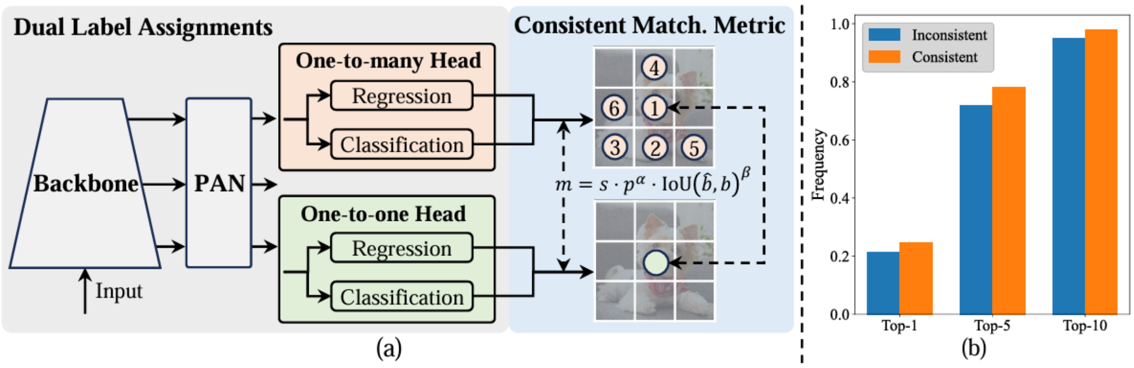


Fig. 9. Architecture of the PGI component (Wang et al., 2024)

4.2. Experiments on Mould Region Detection

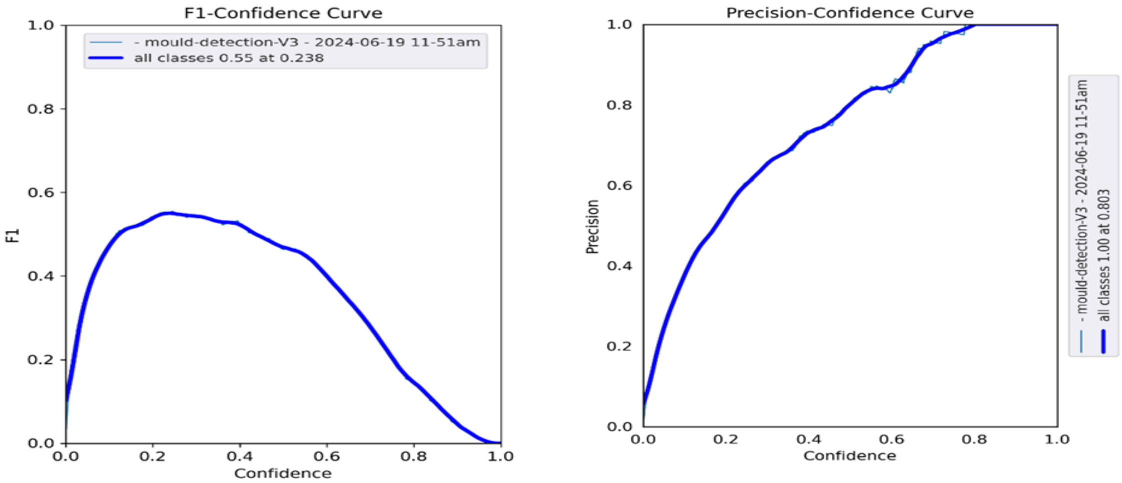
Qualitative results of the proposed model for mould detection are shown in Fig. 10, where one can see from (a) that the baseline of YOLOv9 without adaptation does not fix proper bounding boxes for the mould of different patterns. However, Fig. 10(b) shows the perfect results for moulds with different patterns and background complexities by fixing tight and closed bounding boxes compared to the bounding boxes in Fig. 10(a). This is the impact of the proposed adoption of the YOLOv9 model to detect mould in the images in this work.

(a) Mould region detection before prediction

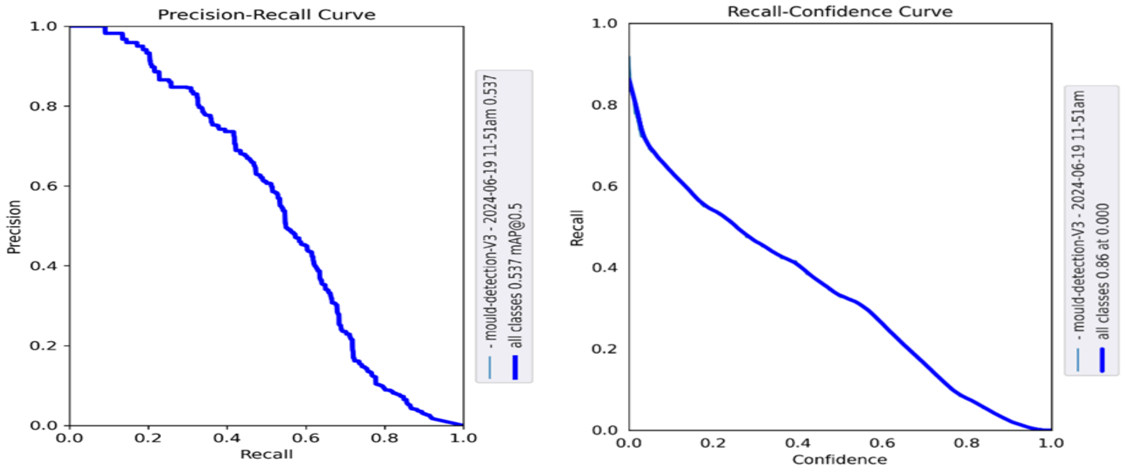
(b) Mould region detection after prediction

Fig. 10. Qualitative mould region detection results of the proposed method before and after prediction

To get a clearer picture from the validation the F1 confidence curve in Fig. 11(a) shows performance value reaching a peak of around 50% and then dropping off. The precision confidence curve also shows a slowing rising precision with maximum confidence of 80%. If we look at the precision-recall curve in Fig. 11(b), we see a balanced precision and recall of just over 50%. The recall confidence curve behaves as expected with recall falling with higher confidence but at 50% confidence, recall is only around 50% which overall is not a great performance. Therefore, one can infer from the analysis that the proposed method focuses on accurate mould detection without any false positives. This is important in the context of real-time applications and practical points of view. Even if the proposed method does not detect the exact boundaries of individual regions of a mould according to the labeled bounding boxes, it still does not affect much as long as it can identify regions of mould in general compared to detecting non-mould regions as a mould. Thus, we can conclude that the proposed system is precise and perfect for mould detection. These conform with the confusion matrix of the validation set as seen in Table 1.



(a) F1 and Precision confidence curve



(b) Precision-recall curve and recall confidence curve

Fig. 11. Performance of the proposed adopted Yolov9 model for the mould region detection

To validate the above statement further, the confusion matrix is estimated by feeding background as input for the mould detection method. The background is defined as the region which falls out of the bounding box in the images. When we feed the background as input to the proposed mould detection, the proposed method reports a 100% classification rate. This indicates, there are no false positives detected by the proposed method. Thus, we can conclude that the proposed method is precise and perfect for mould detection. However, when we consider overall performance including mould to mould and background to background, the average classification rate is 76%, which is competitive and promising if we consider the complexity of the problem. Therefore, there is a scope for improving the proposed method. The performance of the proposed adopted YOLOv9 is compared with the performance of YOLOv10 and found to be better in terms of the average classification rate reported in Table 1. This makes sense because the YOLOv10 performs well for many samples while the proposed model does not. The dataset is new, and the problem is unique and complex, as such there are not many samples that can be generated. Therefore, the proposed work prefers to adopt YOLOv9 to detect mould regions from the images. In the same way, since YOLOv8 is not as advanced as YOLOv9, it does not achieve the best results compared to YOLOv9. However, it is observed from Table 1 that the performance of YOLOv9 is not so different compared to YOLOv8. This shows that the performance of higher versions depends on the number of samples and thus lower versions of YOLO may also be suitable for our work. In general, the YOLOv9 model is simple and effective in feature extraction with a small number of samples and achieves a better average classification rate compared to YOLOv8 and YOLOv10 in this study.

Table 1: Confusion matrix of the proposed adopted YOLOv9 of baseline YOLOv10 models

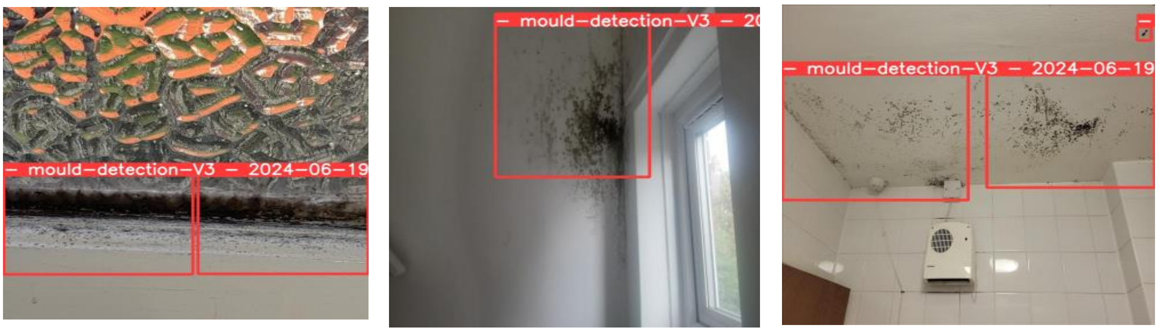
Classes	Proposed adopted YOLOv9		Baseline YOLOv8		Baseline YOLOv10	
	Mould	Background	Mould	Background	Mould	Background
Mould Regions	0.52	0.48	0.50	0.50	0.48	0.52
Background (other than mould regions)	0.00	1.00	0.00	1.00	0.00	1.00
Average Classification Rate	**0.76**		0.75		0.74	

4.3. Discussion

As discussed in the previous section, the accuracy of the proposed method is not very high, but it is promising, and it is precise in detecting moulds of any pattern at any stage. The key reason for the low results is due to improper annotations because the rectangular bounding box does not cover the mould region properly as it has an irregular shape and thus includes background information. This makes learning difficult to extract the effective features which represent mould. One such example can be seen in Fig. 12, where the output of the proposed method does not match with the ground truth. Therefore, the proposed method does not count for calculating the number of overlapping bounding boxes and hence recall is low. This incorrect detection is the result of the irregular pattern of mould region. The proposed method looks for dense patches for accurate mould detection. When the mould regions are scattered and high variations in dense patches, the proposed model does not work well. This is the limitation of the proposed work. To overcome this challenge, it is necessary to annotate at pixel level so that the actual mould region can be used to train the models. This is one more potential area to explore.

This characteristic of mould pattern poses a significant challenge for classification performance metrics like mean average precision (mAp). mAp, by its design, treats each individual bounding box as a distinct target and considers everything outside that box as background. In the context of mould detection, however, this rigid distinction between target and background becomes problematic. As long as the model successfully detects the presence of mould somewhere within the image, the specific location or shape encompassed by the bounding box shouldn't necessarily penalize its performance. What truly matters is the successful identification of mould, regardless of the intricacies of its shape or the slight misalignment of the bounding box

(a) Ground truth of different mould samples



(b) Detection of the proposed model
Fig. 12. Errors in mould region detection

In simpler terms, as long as a detected bounding box captures even a portion of the mould growth within an image, the model has essentially achieved its primary goal. The current metrics, however, might penalize the model for imperfections in bounding box placement, even if the overall objective of mould detection is met. This necessitates the exploration of alternative evaluation metrics that are more tailored to the specific challenges presented by mould detection.

The current focus of the work is to solve the problem of mould detection using Yolov9 architecture but not implementing it in a real-time environment. Therefore, the challenges of real-time implementation are beyond the scope of the proposed work. The key challenge is to make the model robust to external factors, such as non-uniform illumination, outdoor, and indoor, the design of the wall and location effects. In addition, one more key challenge is that the model should be efficient and faster for detecting moulds of any shape and structure. Due to rectangular bounding boxes, the mould region includes outliers and artefacts. This leads to poor performance of the method. These are open challenges for implementing the proposed system in a real-time environment. Since addressing these challenges is beyond the scope of the work, we aim to address them in the future. To overcome these limitations, there is a need to make changes to architecture and add new attention modules to take care of such adverse effects.

Overall, although it looks like the proposed method uses existing YOLOv9 for achieving the results when we count the complexity of the problem, the impact of the solution on society and its significance in solving the current challenge in the UK, the proposed work makes a huge impact and contribution. Therefore, exploring the existing model cannot be regarded as a major weakness of the proposed work

5. Conclusion and Future Work

In this work, we have proposed a newly adopted YOLOv9 for precise mould detection. The key idea of exploring the YOLOv9 lies in its success in object detection in adverse scenarios. This is the first work for mould detection as there are no existing methods developed in the past. We constructed a new dataset along with annotations for experimentation and evaluation. This is another contribution of the proposed work. The proposed work adopts YOLOv9 such that the model considers irregular mould regions as object for mould detection in this work. Exploring the existing model for addressing a complex and unique challenge is the key contribution of the proposed work. Since the proposed model is successful and precise in mould detection, it can make a huge impact on health issues and business. Therefore, the proposed work has great practical implications. The results are compared with the YOLOv10 to show that the YOLOv9 is better for mould detection.

In the present work, the ground truths are rectangular, so the bounding boxes include background information. In addition, the mould is a region, which scatters and grows in arbitrary directions. Therefore, the rectangular bounding box is not the right choice for mould detection. In this situation, if we label pixels through clustering, it is possible to identify the exact region of the mould. This process is likely to improve the precision of the mould detection performance. Thus, pixel annotation is better for labelling the mould regions. This will be considered in our future work.

References

Abdel-Qader, I., Abudayyeh, O., & Kelly, M. E. (2003). Analysis of Edge-Detection Techniques for Crack Identification in Bridges. *Journal of Computing in Civil Engineering*, *17*(4), 255–263. https://doi.org/10.1061/(ASCE)0887-3801(2003)17:4(255)

Annamdas, V. G. M., Bhalla, S., & Soh, C. K. (2017). Applications of structural health monitoring technology in Asia. *Structural Health Monitoring*, *16*(3), 324–346. https://doi.org/10.1177/1475921716653278

Caillaud, D., Leynaert, B., Keirsbulck, M., & Nadif, R. (2018). Indoor mould exposure, asthma and rhinitis: Findings from systematic reviews and recent longitudinal studies. *European Respiratory Review*, *27*(148), 170137. https://doi.org/10.1183/16000617.0137-2017

Damp and mould: Understanding and addressing the health risks for rented housing providers. (2023, September 7). www.gov.uk. Damp and mould: understanding and addressing the health risks for rented housing providers

Davoudi, R., Miller, G. R., & Kutz, J. N. (2018). Structural Load Estimation Using Machine Vision and Surface Crack Patterns for Shear-Critical RC Beams and Slabs. *Journal of Computing in Civil Engineering*, *32*(4), 04018024. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000766

Du, C., Li, B., & Yu, W. (2021). Indoor mould exposure: Characteristics, influences and corresponding associations with built environment—A review. *Journal of Building Engineering*, *35*, 101983. https://doi.org/10.1016/j.jobe.2020.101983

Giakoumoglou, N., Pechlivani, E. M., Sakelliou, A., Klaridopoulos, C., Frangakis, N., & Tzovaras, D. (2023). Deep learning-based multi-spectral identification of grey mould. *Smart Agricultural Technology*, *4*, 100174. https://doi.org/10.1016/j.atech.2023.100174

Hamilton, I., Milner, J., Chalabi, Z., Das, P., Jones, B., Shrubsole, C., Davies, M., & Wilkinson, P. (2015). Health effects of home energy efficiency interventions in England: A modelling study. *BMJ Open*, *5*(4), e007298. https://doi.org/10.1136/bmjopen-2014-007298

Hoang, N.-D. (2018). Image Processing-Based Recognition of Wall Defects Using Machine Learning Approaches and Steerable Filters. *Computational Intelligence and Neuroscience*, *2018*, 1–18. https://doi.org/10.1155/2018/7913952

Jahanshahi, M. R., Jazizadeh, F., Masri, S. F., & Becerik-Gerber, B. (2013). Unsupervised Approach for Autonomous Pavement-Defect Detection and Quantification Using an Inexpensive Depth Sensor. *Journal of Computing in Civil Engineering*, *27*(6), 743–754. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000245

Jones, A., & O'Donoghue, D. (2023, October 17). Awaab Ishak: Family demands "punishment" for boy's mould death. *BBC News*. https://www.bbc.co.uk/news/uk-england-manchester-67424950

Jubayer, F., Soeb, J. A., Mojumder, A. N., Paul, M. K., Barua, P., Kayshar, S., Akter, S. S., Rahman, M., & Islam, A. (2021). Detection of mold on the food surface using YOLOv5. *Current Research in Food Science*, *4*, 724–728. https://doi.org/10.1016/j.crfs.2021.10.003

Koch, C., & Brilakis, I. (2011). Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, *25*(3), 507–515. https://doi.org/10.1016/j.aei.2011.01.002

Kong, Q., Allen, R. M., Kohler, M. D., Heaton, T. H., & Bunn, J. (2018). Structural Health Monitoring of Buildings Using Smartphone Sensors. *Seismological Research Letters*, *89*(2A), 594–602. https://doi.org/10.1785/0220170111

Lee, B. J., & Lee, H. "David." (2004). Position-Invariant Neural Network for Digital Pavement Crack Analysis. *Computer-Aided Civil and Infrastructure Engineering*, *19*(2), 105–118. https://doi.org/10.1111/j.1467-8667.2004.00341.x

Li, S and Calderon, A. D. (2023). Review of the current research status and constraints of wall-climbing robots. *In Proc. International Conference on Robotics, Control and Automation Engineering*, pp 44-49.

Lorenzoni, F., Casarin, F., Caldon, M., Islami, K., & Modena, C. (2016). Uncertainty quantification in structural health monitoring: Applications on cultural heritage buildings. *Mechanical Systems and Signal Processing*, *66–67*, 268–281. https://doi.org/10.1016/j.ymssp.2015.04.032

Menneer, T., Mueller, M., Sharpe, R. A., & Townley, S. (2022). Modelling mould growth in domestic environments using relative humidity and temperature. *Building and Environment*, *208*, 108583. https://doi.org/10.1016/j.buildenv.2021.108583

Mimura, T., & Mita, A. (2017). Automatic Estimation of Natural Frequencies and Damping Ratios of Building Structures. *Procedia Engineering*, *188*, 163–169. https://doi.org/10.1016/j.proeng.2017.04.470

Mita, A. (2016). *Gap between technically accurate information and socially appropriate information for structural health monitoring system installed into tall buildings* (T. Kundu, Ed.; p. 98050E). https://doi.org/10.1117/12.2219065

Nair, A., Manomohan, N. K., Reddy, H., Pandey, S., & Kadam, P. (2021). Fungus Detection and Identification using Computer Vision Techniques and Convolution Neural Networks. *2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, 1–6. https://doi.org/10.1109/SMARTGENCON51891.2021.9645827

Noel, A. B., Abdaoui, A., Elfouly, T., Ahmed, M. H., Badawy, A., & Shehata, M. S. (2017). Structural Health Monitoring Using Wireless Sensor Networks: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, *19*(3), 1403–1423. https://doi.org/10.1109/COMST.2017.2691551

Oh, B. K., Kim, K. J., Kim, Y., Park, H. S., & Adeli, H. (2017). Evolutionary learning based sustainable strain sensing model for structural health monitoring of high-rise buildings. *Applied Soft Computing*, *58*, 576–585. https://doi.org/10.1016/j.asoc.2017.05.029

Pahlberg, T., Thurley, M., Popovic, D., & Hagman, O. (2018). Crack detection in oak flooring lamellae using ultrasound-excited thermography. *Infrared Physics & Technology*, *88*, 57–69. https://doi.org/10.1016/j.infrared.2017.11.007

Perez, H., Tah, J. H. M., & Mosavi, A. (2019). Deep Learning for Detecting Building Defects Using Convolutional Neural Networks. *Sensors*, *19*(16), 3556. https://doi.org/10.3390/s19163556

Pragalath, H., Seshathiri, S., Rathod, H., Esakki, B., & Gupta, R. (2018). Deterioration Assessment of Infrastructure Using Fuzzy Logic and Image Processing Algorithm. *Journal of Performance of Constructed Facilities*, *32*(2), 04018009. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001151

Radopoulou, S. C., & Brilakis, I. (2017). Automated Detection of Multiple Pavement Defects. *Journal of Computing in Civil Engineering*, *31*(2), 04016057. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000623

Sharpe, R. A., Cocq, K. L., Nikolaou, V., Osborne, N. J., & Thornton, C. R. (2016). Identifying risk factors for exposure to culturable allergenic moulds in energy efficient homes by using highly specific monoclonal antibodies. *Environmental Research*, *144*, 32–42. https://doi.org/10.1016/j.envres.2015.10.029

Simon-Nobbe, B., Denk, U., Pöll, V., Rid, R., & Breitenbach, M. (2008). The Spectrum of Fungal Allergy. *International Archives of Allergy and Immunology*, *145*(1), 58–86. https://doi.org/10.1159/000107578

Soeb, Md. J. A., Jubayer, Md. F., Tarin, T. A., Al Mamun, M. R., Ruhad, F. M., Parven, A., Mubarak, N. M., Karri, S. L., & Meftaul, I. Md. (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). *Scientific Reports*, *13*(1), 6078. https://doi.org/10.1038/s41598-023-33270-4

Sohan, M., Ram, T. S and Reddy, V. R., (2023), A review on YOLOv8 and its advancements, In Proc. Data Intelligence and Cognitive Informatics, pp 529-545.

Song, G., Wang, C., & Wang, B. (2017). Structural Health Monitoring (SHM) of Civil Structures. *Applied Sciences*, *7*(8), 789. https://doi.org/10.3390/app7080789

Valero, E., Forster, A., Bosché, F., Hyslop, E., Wilson, L., & Turmel, A. (2019). Automated defect detection and classification in ashlar masonry walls using machine learning. *Automation in Construction*, *106*, 102846. https://doi.org/10.1016/j.autcon.2019.102846

Valero, E., Forster, A., Bosché, F., Renier, C., Hyslop, E., & Wilson, L. (2018, July 22). *High Level-of-Detail BIM and Machine Learning for Automated Masonry Wall Defect Surveying*. 34th International Symposium on Automation and Robotics in Construction, Taipei, Taiwan. https://doi.org/10.22260/ISARC2018/0101

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). *YOLOv10: Real-Time End-to-End Object Detection* (arXiv:2405.14458). arXiv. http://arxiv.org/abs/2405.14458

Wang, C.-Y., Yeh, I.-H., & Liao, H.-Y. M. (2024). *YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information* (arXiv:2402.13616). arXiv. http://arxiv.org/abs/2402.13616

Zhang, F., Yang, Y., Xiong, H., Yang, J., & Yu, Z. (2019). Structural health monitoring of a 250-m super-tall building and operational modal analysis using the fast Bayesian FFT method. *Structural Control and Health Monitoring*, e2383. https://doi.org/10.1002/stc.2383