REAL-TIME DETECTION OF CRACKS AND MOISTURE DEFECTS IN BUILDINGS: A Two-Part Report on Supervised and Unsupervised Methods

B. Daribayev [1], A. Mukhanbet [1]†, Y. Kenzhebek [1], N. Kassymbek [1], Zh. Buribayev [1], T. Imankulov [1] and I. Starikova [3], M. Zyskin [4]‡ and A. Stone [2]

[1] *Al-Farabi Kazakh National University, Almaty, Kazakhstan*
[2] *Surveyr Company*
[3] *Slovak Academy of Sciences*
[4] *Consultant*

(*Communicated to* MIIR *on 5 September 2025*)

Study Group: 187th European Study Group with Industry (ESGI 187), 14–18 July 2025, University of Bristol, UK

Communicated by: Zahraa S. Abdallah, Matthew G. Hennessy, and Stuart J. Thomson, University of Bristol

Industrial Partner: Surveyr Company

Presenter: Amy Stone

Team Members: Zh. Buribayev, Al-Farabi Kazakh National University;
B. Daribayev, Al-Farabi Kazakh National University;
T. Imankulov, Al-Farabi Kazakh National University;
N. Kassymbek, Al-Farabi Kazakh National University;
Y. Kenzhebek, Al-Farabi Kazakh National University;
A. Mukhanbet, Al-Farabi Kazakh National University;
I. Starikova, Slovak Academy of Sciences;
M. Zyskin, consultant;
A. Stone, Surveyr Company.

Industrial Sector: Construction

Key Words: structural defect detection; computer vision; building inspection; crack analysis; moisture detection

MSC2020 Codes: 68T45; 68T07

† Corresponding author (Part 1): <mukhanbetaksultan0414@gmail.com>
‡ Corresponding author (Part 2): maximzyskin@gmail.com

Summary

Building inspection and defect detection remain critical challenges in building surveying, traditionally requiring extensive expertise and time. This paper presents a solution for a problem posed by Surveyr for real-time detection, classification, and analysis of structural defects, focusing on cracks and moisture-related damages. The system combines custom Convolutional Neural Networks (CNN), ResNet-50, and YOLO v11 to address different analytical tasks. Two specialized datasets were used: a crack dataset of 518 annotated images classified into five width categories per BRE Digest 251 standards, and a moisture dataset of 245 images in three types per BRE Digest 245 guidelines. Preprocessing employed manual annotation (LabelImg, Roboflow) and automated filtering with a 40,000-image pre-labeled dataset. The custom CNN has 5.4M trainable parameters, optimized for 227×227×3 RGB inputs using binary cross-entropy loss and Adam optimization. ResNet-50 was applied for multi-class crack classification, while YOLO v11 handled integrated detection and classification. Large-scale images were processed via a sliding window, with confidence-based outputs for field use. Deployment includes a mobile app for on-site assessment and a web platform for detailed analysis. The system shows strong performance in both crack and moisture detection, integrating multiple deep learning architectures for practical inspection workflows. Future work will expand datasets, integrate Large Language Models for automated reports, and employ stereo cameras for precise crack width measurement, advancing automated building inspection in line with industry standards.

1 Introduction

1.1 Background and Motivation

The structural integrity of buildings and civil infrastructure represents a critical concern for public safety, economic sustainability, and urban development. Traditional inspection methods for structural defects rely heavily on manual assessment by trained professionals, a process that is inherently time-consuming, labor-intensive, and susceptible to human error and subjectivity [1]. The emergence of artificial intelligence (AI) and computer vision technologies has opened new avenues for automated defect detection in civil infrastructure, offering the potential for more consistent, efficient, and objective assessment methodologies.

Structural defects, particularly cracks and moisture-related damage, are among the most common indicators of building deterioration. Cracks can develop due to various factors including structural loading, thermal expansion and contraction, foundation settlement, and material aging [2]. Similarly, moisture infiltration can lead to serious structural problems including material degradation, mold growth, and compromised building

envelope performance. The early detection and accurate classification of these defects are essential for maintaining building safety and preventing costly remedial work.

Current manual inspection processes for building defects face several fundamental limitations. First, the detection and classification of structural anomalies require specialized expertise and extensive field experience. Second, the scale and complexity of modern buildings make comprehensive manual inspection economically prohibitive and practically challenging. Third, the subjective nature of visual assessment leads to inconsistencies in defect classification and severity evaluation. Fourth, traditional methods often lack the capability for real-time analysis and immediate documentation of findings.

The specific challenges addressed in this work include: (1) accurate detection of crack presence in building surfaces under varying lighting and environmental conditions; (2) precise classification of crack types and severity levels according to established industry standards; (3) identification and categorization of moisture-related defects; and (4) development of practical mobile solutions for field deployment that can operate reliably in real-world conditions.

The application of deep learning techniques to crack detection has evolved significantly in recent years. Dorafshan et al. (2018) conducted one of the seminal comparative studies between deep convolutional neural networks and traditional edge detection methods for crack detection in concrete, demonstrating that CNN-based approaches significantly outperformed conventional image processing techniques [3]. This work established the foundation for subsequent research in AI-driven crack detection systems.

Recent advances have demonstrated remarkable improvements in detection accuracy. A comprehensive comparative analysis of deep learning models for crack detection in buildings achieved accuracy rates exceeding 99.98% using models such as Inception V3 and ResNet-50 [4]. The study evaluated multiple architectures and demonstrated the superiority of deep learning approaches over traditional computer vision methods for crack detection applications. Zhang et al. (2024) developed an advanced road crack detection system using deep convolutional neural networks, focusing on segmentation-based approaches that can precisely delineate crack boundaries [5]. Their work highlighted the importance of architectural choices in achieving robust performance across diverse environmental conditions and crack characteristics.

The evolution of neural network architectures has significantly impacted crack detection performance. Deep neural networks specifically designed for crack detection inside structures have shown exceptional capability in identifying internal structural defects using DenseNet architectures [6]. This research demonstrated that deeper networks with dense connectivity patterns can capture subtle crack features more effectively than traditional CNN approaches.

Transfer learning has emerged as a particularly effective strategy for crack detection applications, where pre-trained models are fine-tuned on domain-specific datasets. This approach has proven especially valuable given the typically limited size of crack detection datasets compared to general computer vision applications [7].

The You Only Look Once (YOLO) family of algorithms has gained prominence in real-time object detection applications. Recent iterations, particularly YOLO v8 and v11, have shown remarkable performance in detecting small objects and fine-grained features, making them suitable for crack detection applications [8, 9]. The single-stage

detection approach of YOLO architectures offers computational efficiency advantages for mobile deployment scenarios.

Integration of attention mechanisms and multi-scale feature fusion has further enhanced YOLO performance for crack detection. Li et al. (2023) developed a bridge crack detection system based on SSC-YOLO with multi-scale feature fusion, achieving improved detection accuracy for cracks of varying sizes [10]. This work demonstrated the effectiveness of combining attention mechanisms with YOLO architectures for infrastructure inspection applications.

While crack detection has received substantial attention in the literature, moisture-related defect detection remains less explored in computer vision applications. Traditional moisture detection methods rely on thermal imaging, electrical resistance measurements, and visual inspection [11, 12]. However, recent developments in computer vision and deep learning have shown promise for automated moisture detection using conventional RGB imagery. The challenge of moisture detection lies in the subtle visual indicators that may not be immediately apparent in standard photographic imagery. Research in this area has focused on developing specialized feature extraction techniques that can identify visual patterns associated with different types of moisture damage, including rising damp, penetrating damp, and condensation-related issues [13].

The trend towards integrated systems that can detect multiple types of defects simultaneously has emerged as a significant research direction. These systems typically combine multiple neural network architectures or employ ensemble methods to achieve comprehensive defect analysis [14]. The advantage of integrated approaches lies in their ability to provide holistic building assessment capabilities within a single analytical framework.

Recent research has explored the development of unified frameworks for detecting multiple types of structural defects, including cracks, spalling, and corrosion, using single deep learning models [15]. Such approaches offer practical advantages for field deployment by reducing the computational overhead and system complexity associated with multiple specialized models.

The practical deployment of AI-based defect detection systems on mobile platforms presents unique challenges related to computational efficiency, user interface design, and real-world performance. Mobile applications require optimized models that can operate within the computational constraints of smartphone and tablet devices while maintaining acceptable accuracy levels [16].

Recent advances in mobile computing and edge AI have made sophisticated computer vision applications increasingly feasible for field deployment. Model optimization techniques such as quantization, pruning, and knowledge distillation have enabled the deployment of complex deep learning models on resource-constrained mobile devices [17, 18].

The availability of high-quality, standardized datasets has been crucial for advancing crack detection research. The concrete crack images dataset developed by Özgenel (2019) has become a benchmark resource, containing 40,000 images for binary crack classification [19]. This dataset has enabled comparative studies and facilitated the development of robust crack detection algorithms.

However, the field still lacks comprehensive datasets that align with established industry classification standards such as those provided by the Building Research Establishment (BRE) Digest series. The BRE Digest 251 provides standardized guidelines for

assessing damage in low-rise buildings, while BRE Digest 245 addresses rising damp diagnosis and treatment [20, 21]. Integration of these standards into AI-based detection systems remains an underexplored area.

Despite significant progress in AI-based defect detection, several gaps remain in the current literature. First, most studies focus exclusively on crack detection, with limited attention to other common building defects such as moisture damage. Second, few studies have developed integrated systems capable of detecting and classifying multiple defect types within a single framework using industry-standard classification schemes. Third, the transition from research prototypes to practical field applications remains challenging, with limited studies addressing real-world deployment considerations such as varying lighting conditions, diverse building materials, and user interface requirements. Fourth, there is insufficient research on the integration of multiple deep learning architectures within unified systems that can leverage the strengths of different approaches.

This work aims to address the identified research gaps by developing a comprehensive AI-driven system for building defect detection that encompasses both crack and moisture-related damage. The specific objectives include:

(1)Development of accurate detection algorithms for crack presence using custom CNN architectures optimized for building inspection scenarios

(2)Implementation of multi-class classification systems for crack type and severity assessment according to BRE Digest 251 standards

(3)Creation of specialized models for moisture defect detection and classification based on BRE Digest 245 guidelines

(4)Integration of multiple deep learning architectures (CNN, ResNet, YOLO) within a unified analytical framework

(5)Development of practical mobile and web-based applications for field deployment

(6)Evaluation of system performance across diverse real-world building inspection scenarios

2 Methodology

2.1 Overall System Architecture

The defect detection system developed for *Surveyr* is a multi-level process for automated real-time detection and classification of structural defects. The architecture combines three key deep learning modules, each responsible for a specific stage of analysis:

- A custom CNN module for binary detection of defect presence.
- ResNet-50 for detailed multi-class crack classification.
- YOLO v11 for simultaneous detection and classification of defects in full-frame images.

2.2 Datasets

The system was trained and tested on two specialized datasets:

- **Crack Dataset** — 518 annotated crack images classified into five width categories according to BRE Digest 251: hairline cracks, 1–5 mm, 5–15 mm, 15–25 mm, and ¿25 mm.

•**Moisture Dataset** — 245 images of damage categorized into three types according to BRE Digest 245: rising damp, penetrating damp, and condensation.
To enhance feature variability, an additional dataset of 40,000 pre-labeled images was used for automatic filtering and model pre-training [22].

2.2.1 *Data Preprocessing for Crack Analysis*

To ensure high accuracy in defect detection and classification within the defect detection system developed for *Surveyr*, a two-stage data preprocessing strategy was developed, incorporating both manual annotation and automated filtering based on pre-trained models.

2.2.2 *Manual Annotation*

At the first stage, specialized tools LabelImg and Roboflow were used to create annotations in the YOLO format. Defects were manually delineated over crack regions in the original images, ensuring high-quality labeling and compliance with BRE Digest 251 standards.

Following annotation, crack region cropping was applied, allowing irrelevant image context to be excluded from analysis and enabling the models to focus on the textural and structural features of the damage. To improve the robustness of the models against variations in shooting conditions and defect shapes, data augmentation techniques were applied, including:

•Geometric transformations ($\pm 90^\circ$ rotations, vertical and horizontal flips),
•Photometric adjustments (brightness and contrast modification),
•Artificial noise injection to simulate degraded image quality in field conditions.

2.2.3 *Model-Oriented Preprocessing*

At the second stage, automated image filtering was applied using a pre-trained convolutional architecture (ResNet). As input, a large dataset of 40,000 images annotated with a binary label (crack presence/absence) was used. The workflow included the following steps:

•Prediction of crack presence in each image using a ResNet/EfficientNet model.
•Filtering: images with a low probability of containing a crack were excluded from further processing.
•Forwarding the filtered images to the augmentation pipeline, analogous to Method 1, for preparing the data for training the main model.
Here, a targeted data augmentation strategy was used to improve the robustness of the models to a variety of shooting conditions and defect variability. The process involved a set of geometric and photometric transformations aimed at simulating real-world building inspection scenarios. Techniques used included random rotations and mirror reflections to ensure invariance to defect orientation; scaling and random cropping to allow the models to effectively handle cracks of varying sizes; changes in brightness, contrast, and saturation to compensate for differences in lighting and color rendition between cameras;

and the addition of artificial Gaussian noise to simulate interference that occurs when shooting in the field. Figure 1 shows how the augmentation techniques are used for cracks.

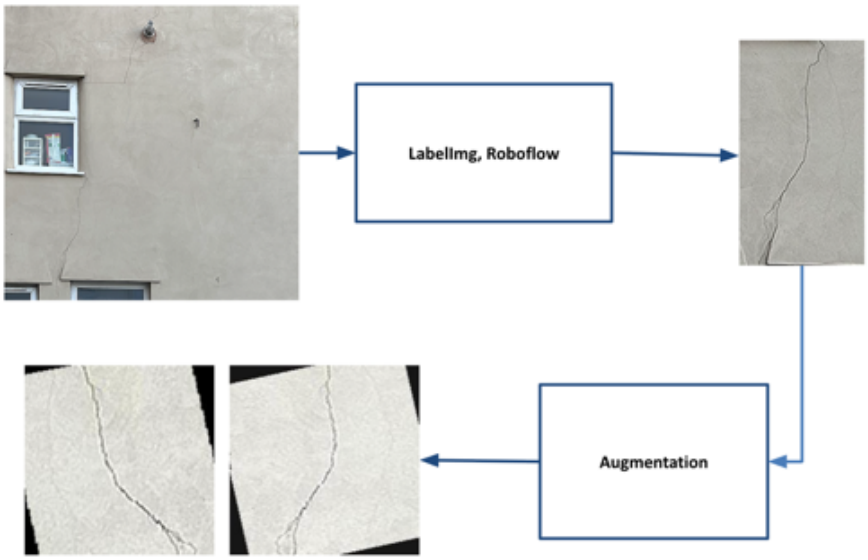

Figure 1. Crack augmentation techniques.

One of the approaches used in the defect detection system was the use of the Roboflow platform for automated augmentation and dataset management. As shown in Figure 2, Roboflow allowed not only to centrally store and version datasets, but also to apply built-in augmentation algorithms in pipeline mode. The transformations used included automatic normalization and resizing of images to meet the input requirements of the models, random rotations and reflections, brightness and contrast variations, and the imposition of synthetic noise.

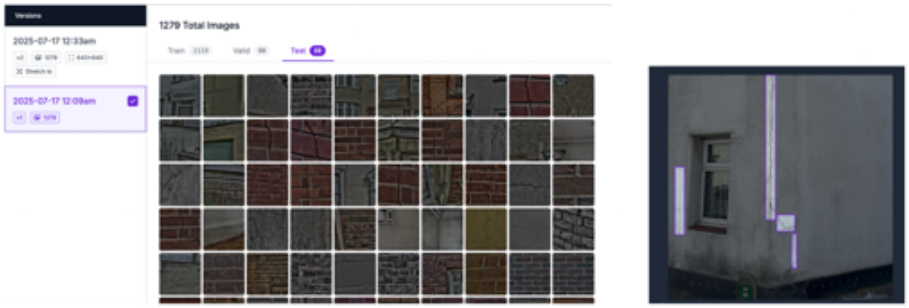

Figure 2. Data augmentation in Roboflow.

For the crack detection task, an adapted convolutional neural network architecture was used, focused on binary classification (presence/absence of a crack). The model accepted

color images of $227 \times 227 \times 3$ (RGB) as input, while all 5.4 million parameters were trainable, which provided flexibility in adjusting to the specifics of construction defects.

The ResNet-50 architecture was used to solve the problem of multi-class crack width classification according to the BRE Digest 251 standard. The architecture of the ResNet-50 neural network is shown in Figure 3. The model was chosen due to its deep residual structure, which effectively solves the problem of vanishing gradients and allows extracting high-level features even from complex texture images of damage.

Layer (type)	Output Shape	Param #
conv2d_7 (Conv2D)	(None, 223, 223, 6)	456
activation_16 (Activation)	(None, 223, 223, 6)	0
average_pooling2d_7 (Average	(None, 111, 111, 6)	0
conv2d_8 (Conv2D)	(None, 107, 107, 16)	2416
activation_17 (Activation)	(None, 107, 107, 16)	0
average_pooling2d_8 (Average	(None, 53, 53, 16)	0
flatten_4 (Flatten)	(None, 44944)	0
dense_10 (Dense)	(None, 120)	5393400
activation_18 (Activation)	(None, 120)	0
dense_11 (Dense)	(None, 84)	10164
activation_19 (Activation)	(None, 84)	0

...
Total params: 5,406,521
Trainable params: 5,406,521
Non-trainable params: 0

Figure 3. Neural network architecture (ResNet-50).

The input data were RGB images, pre-scaled and normalized. Transfer learning was used for training: the model was initialized with weights pre-trained on ImageNet, after which the final fully connected layers were replaced with a classifier adapted to five classes of cracks (fine cracks, 1–5 mm, 5–15 mm, 15–25 mm, ¿25 mm). The training process included the use of the Categorical Crossentropy loss function and the Adam optimizer, which ensured a balance between convergence speed and stability. Accuracy, Precision, Recall and F1-score were used as key metrics for a more accurate assessment of the classification quality of each defect type.

3 Results

The proposed system demonstrated reliable performance in detecting and classifying cracks on various structural surfaces. Figure 4 presents the results of binary classification, where the system successfully and with high accuracy distinguished between concrete surfaces with and without cracks. The CNN-based classifier effectively identified even barely visible cracks, including hairline fractures and more pronounced structural defects, while correctly rejecting false positives associated with surface texture, shadows, and common concrete imperfections. The classification remained stable under different lighting conditions and surface textures: the system consistently labeled samples with visible cracks as "Crack" and surfaces without cracks as "No Crack."

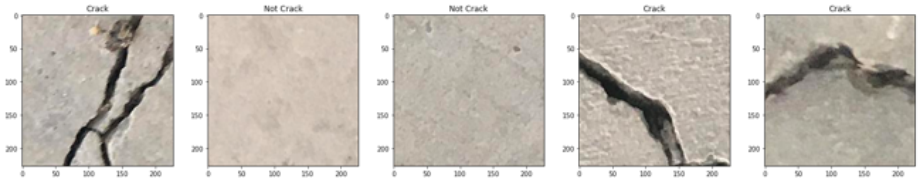

Figure 4. Binary classification results with ResNet.

The results of crack classification and segmentation were based on a data set of 40,000 data points. The effectiveness of the segmentation is demonstrated in Figure 5, which shows the system's ability to accurately localize and delineate crack boundaries in larger structural images. The raw image processing pipeline successfully identified the crack path along the concrete surface, with the highlighted segmentation mask (shown in pink overlay) accurately tracking the crack path from initiation to termination.

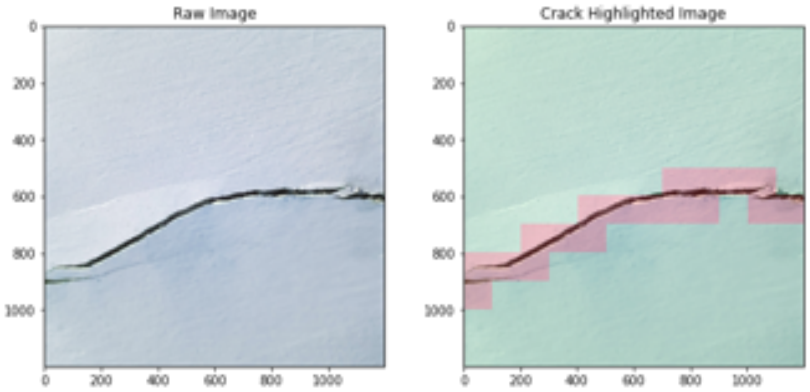

Figure 5. Crack segmentation results.

The segmentation algorithm effectively processed the geometry of curved cracks while preserving spatial accuracy, demonstrating the system's capability to provide detailed geometric information necessary for structural assessment. Its confident region extraction approach ensures that only areas with a high probability of crack presence are highlighted, thereby reducing the number of false positives on complex surfaces. The ResNet-50 model

Table 1. Results of different metrics for the ResNet-50 model.

Metrics	**Categories 0,1&2 (Aesthetic)**	**Categories 3&4 (Serviceability)**	**Categories 5 (Stability)**
Accuracy	0.72	0.72	0.72
Precision	0.80	0.65	1.00
Recall	0.78	0.80	0.22
F1-score	0.79	0.72	0.36

achieved an overall accuracy of 72% across the three crack severity categories defined by BRE Digest 251. Performance metrics revealed distinct patterns for different crack classifications. For categories 0, 1, and 2 (aesthetic cracks), the model demonstrated balanced performance with a precision of 0.80 and a recall of 0.78, resulting in an F1-score of 0.79 for 36 test samples. This indicates a high capability in identifying minor aesthetic cracks, which pose minimal structural concerns. Table 1 presents all metrics and results of the ResNet-50 model.

Categories 3 and 4 (serviceability-related cracks) demonstrated moderate precision of 0.65 but higher recall of 0.80, resulting in an F1-score of 0.72 across 35 test samples. The lower precision suggests the presence of some false positives, where aesthetic cracks were mistakenly classified as serviceability issues. Category 5 (stability-related cracks) showed perfect precision of 1.00 but considerably lower recall of 0.22, leading to an F1-score of 0.36 across 9 samples. This pattern indicates a conservative approach by the model toward critical structural defects, avoiding false alarms but potentially missing some severe cases.

The confusion matrix provides detailed insight into classification patterns and the distribution of errors across crack severity categories. For categories 0, 1, and 2 (aesthetic cracks), the model correctly classified 28 out of 36 samples, while 8 samples were misclassified as serviceability-related cracks, and no misclassifications occurred with stability-related cracks. This 77.8% accuracy in the aesthetic category demonstrates reliable detection of minor surface defects. Categories 3 and 4 (serviceability cracks) yielded 28 correct classifications out of 35 samples, with 7 samples misclassified as aesthetic cracks, and no false positives observed for stability-related issues. This 80% within-category accuracy highlights the model's solid distinction between minor and moderate crack severities. Figure 6 presents the confusion matrix for ResNet-50.

Category 5 (stability-related cracks) represented the most challenging classification scenario: only 2 out of 9 samples were correctly identified, while 7 samples were misclassified as serviceability-related cracks. A parallel comparison of annotated images (left) and model predictions (right), presented in Figure 7, demonstrates the practical effectiveness of the ResNet-50 classifier across different crack scenarios and construction materials. The test dataset covered a variety of surface types, including brickwork, concrete, stone, and painted surfaces, reflecting realistic building inspection conditions.

Visual inspection reveals overall stable classification performance across different materials and lighting conditions. The model successfully identified crack patterns both in controlled laboratory settings and in field conditions, maintaining classification accuracy

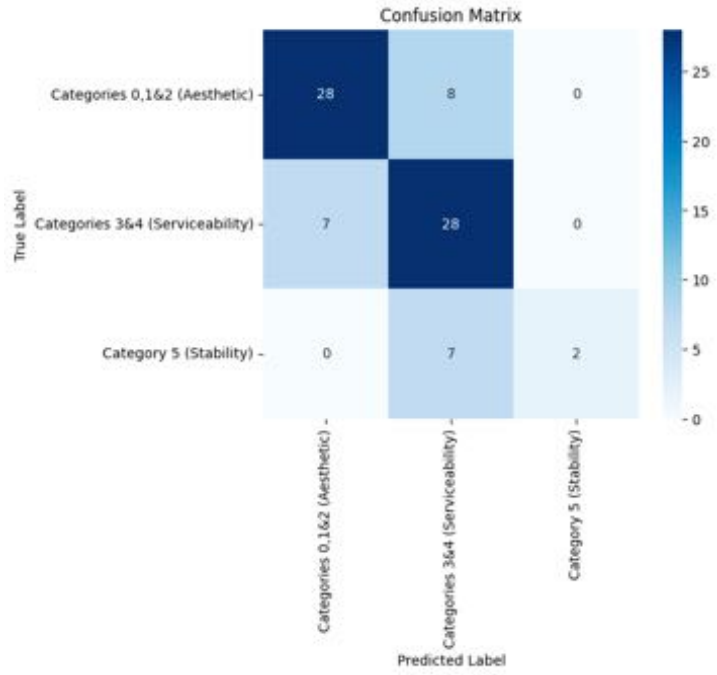

Figure 6. Confusion matrix for ResNet-50.

under varying image quality and crack characteristics. Noticeable performance fluctuations were observed in cases involving complex surface textures, shadows, or weathering effects, which can potentially affect crack visibility.

Figure 7. Side-by-side comparison of labeled images (left) and model predictions (right).

The classification color boundaries (shown in cyan, red, and blue frames) indicate the model's confidence in assigning categories, with a clear separation between classifications of aesthetics, serviceability, and stability. In some cases, borderline situations were

observed, where visual similarities between adjacent severity categories posed challenges both for human annotators and for the automated system—particularly in distinguishing between moderate cracks affecting serviceability and severe defects threatening structural stability.

The YOLO v11 model demonstrated excellent convergence during the 100-epoch training process across multiple performance metrics, as shown in Figure 8. The training loss curve decreased smoothly and exponentially from an initial value of approximately 0.9 to a final convergence around 0.05, indicating efficient learning without significant fluctuations or instability. Validation loss followed a similar trend, dropping from 1.2 to approximately 0.35, with the smoothed curve showing steady improvement throughout the training process. The top-1 accuracy metric consistently improved from an initial value of 0.65 to a final value of 0.9, reflecting the model's increasing ability to correctly classify the primary crack category.

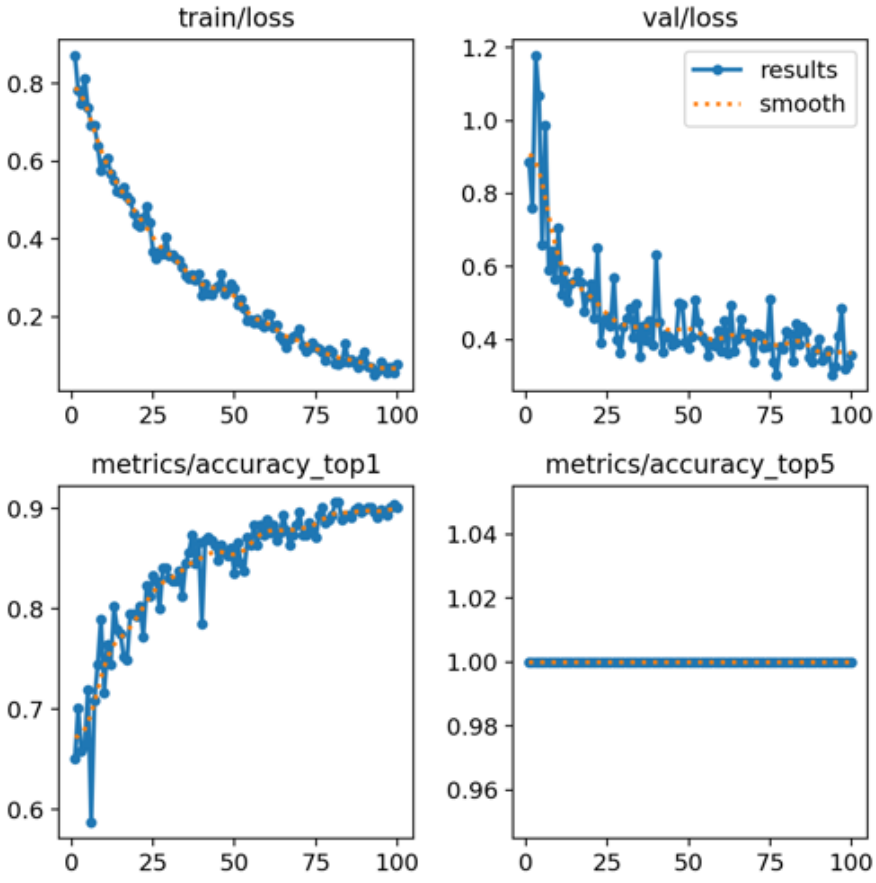

Figure 8. Metrics results for YOLO v11.

In particular, the top-5 accuracy consistently maintained a perfect score of 1.0 throughout the entire training process, indicating that the correct classification was always in-

cluded among the model's top five predictions. The absence of significant overfitting, as evidenced by the parallel dynamics of training and validation loss curves, highlights the model's strong generalization capability.

The YOLO v11 confusion matrix, presented in Figure 9, demonstrates outstanding performance in multi-class crack classification with minimal inter-category confusion. The aesthetic crack category achieved remarkable accuracy: 163 correct classifications out of 183 samples (89.1%), with only 19 cases misclassified as serviceability-related and a single false positive for stability. The serviceability crack category showed similarly robust performance: 101 correct predictions out of 116 samples (87.1%), including 5 misclassified as aesthetic and 10 as stability, indicating some expected overlap between adjacent severity levels.

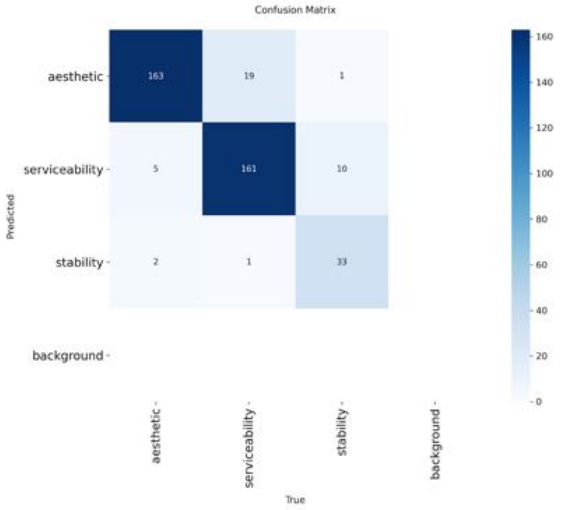

Figure 9. YOLO v11 Error Matrix.

The stability category demonstrated strong discrimination: 33 correct classifications out of 36 samples (91.7%), with minimal confusion—only 2 cases were mistakenly classified as aesthetic and 1 as serviceability. The background class appeared to be perfectly classified without false positives or false negatives, although the matrix shows an empty row, which may indicate either perfect classification or a formatting artifact in the data presentation.

The comparison between ground-truth labeled objects and YOLO v11 detections, presented in Figure 10, illustrates the model's comprehensive object detection capabilities across various structural scenarios. The detection system successfully identified and localized cracks while simultaneously providing accurate severity classification, as evidenced by bounding boxes with confidence scores and category labels. The model demonstrated reliable performance across different surface materials, including concrete, brick, stone, and plastered surfaces, maintaining consistent detection accuracy regardless of texture complexity or lighting conditions. Notable advantages include precise localization of both

prominent structural cracks and barely visible surface defects, with bounding boxes accurately enclosing the corresponding crack regions.

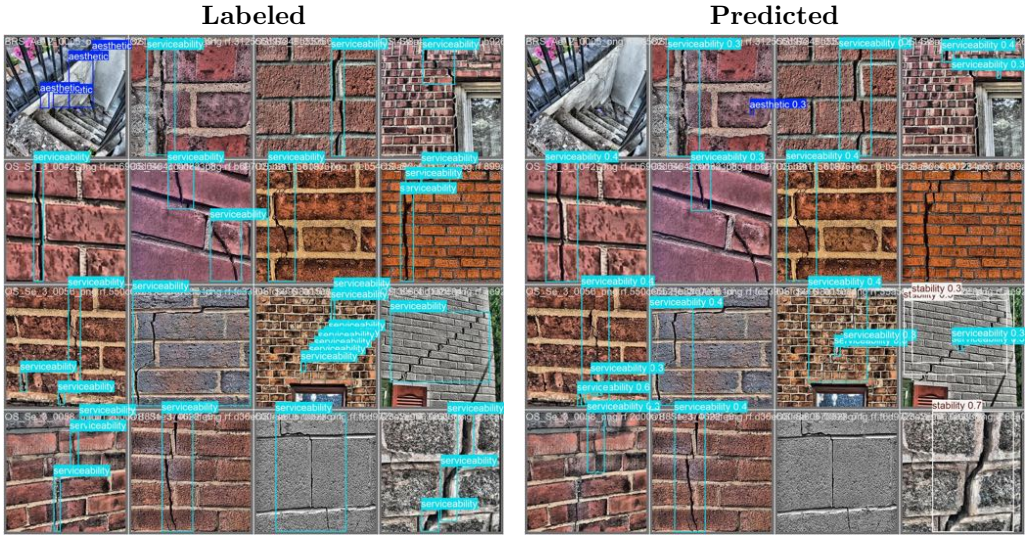

Figure 10. Comparison of the results of detection of labeled true objects and YOLO v11.

The confidence scores displayed alongside each detection provide a quantitative measure of prediction reliability, enabling field inspectors to prioritize findings based on both severity classification and the model's confidence level. In certain cases, the model demonstrated the ability to detect multiple crack instances within a single image while maintaining individual classification accuracy for each defect. Detection results showed a minimal number of false positives, with the system effectively distinguishing true structural defects from surface artifacts such as mortar joints, shadows, or superficial stains. Color-coded bounding boxes corresponding to different severity categories (aesthetic, serviceability, stability) provide instant visual feedback for rapid assessment during field inspections, highlighting the practical value of the integrated detection and classification approach for real-world building inspection workflows.

The crack classification system demonstrated high effectiveness across various structural surfaces and crack morphologies in a comparative analysis between ground-truth annotations and model predictions. The results are presented in Figure 11. The visual outputs illustrate the system's ability to accurately identify and classify different types of cracks across materials such as concrete, brick, stone, and finished surfaces. A notable performance example (highlighted with a red rectangle) shows a complex branching crack pattern in concrete that was correctly identified and classified, demonstrating the model's capability to handle nonlinear crack geometries. The classification system maintained consistency across varying surface textures, lighting conditions, and crack widths, successfully distinguishing fine aesthetic cracks from more pronounced structural defects.

The model demonstrated particular effectiveness in dealing with complex scenarios such as cracks intersecting surface joints, weather conditions and colour variations in

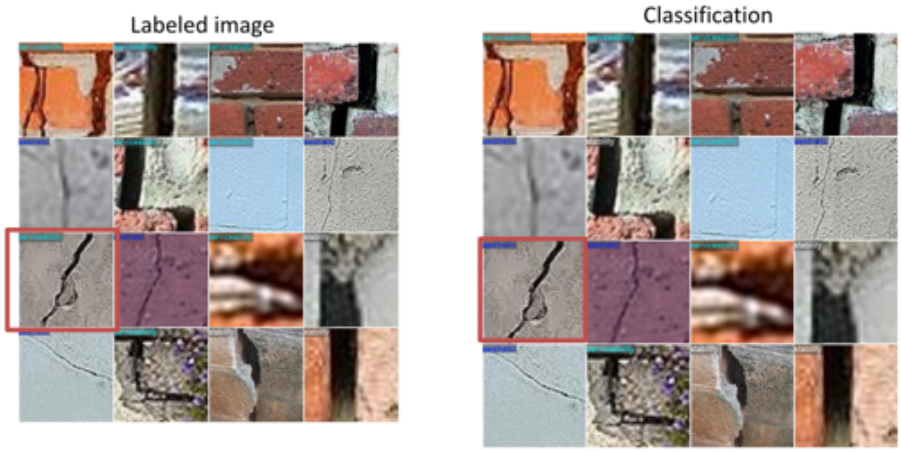

Figure 11. Comparative analysis between actual data obtained using labeling and model predictions.

building materials. Some examples demonstrate the conservative approach of the system, where borderline cases between severity categories were classified with due caution.

3.1 Moisture classification results

The moisture detection model demonstrated fast convergence and stable learning dynamics. In Figure 12, it can be seen that the training loss dropped sharply from an initial value of 1.0 to approximately 0.05 over the first 10 epochs, followed by a gradual decrease to values close to zero, indicating effective feature learning for moisture pattern recognition. The validation loss followed a similar trend, dropping from 1.15 to approximately 0.08, with the smoothed curve showing consistent improvement without significant overfitting. The top-1 accuracy metric showed impressive progression over training, starting at 0.45 and quickly peaking at 0.98 by the 20th epoch, maintaining this high performance throughout the remainder of the training period.

The top-5 accuracy achieved perfect classification (1.0) from the early epochs and maintained these results, indicating that the correct humidity classification consistently appeared among the model's most reliable predictions. The parallel convergence of training and validation metrics suggests an optimal balance of the dataset and an appropriate model complexity for the task of three-class humidity classification.

The confusion matrix for humidity classification (Figure 13) demonstrates the exceptional ability to distinguish between the three categories of moisture damage defined by BRE Digest 245. For condensation, perfect classification accuracy was achieved: 37 out of 37 samples were correctly identified, with no misclassifications. This result highlights the model's capability to recognize characteristic changes in surface moisture, thermal effects, and damage associated with evaporation. Penetrating dampness showed nearly perfect performance, with 41 correct classifications out of 42 samples (97.6% accuracy), and only one instance misclassified as rising damp, indicating minimal confusion between water ingress mechanisms.

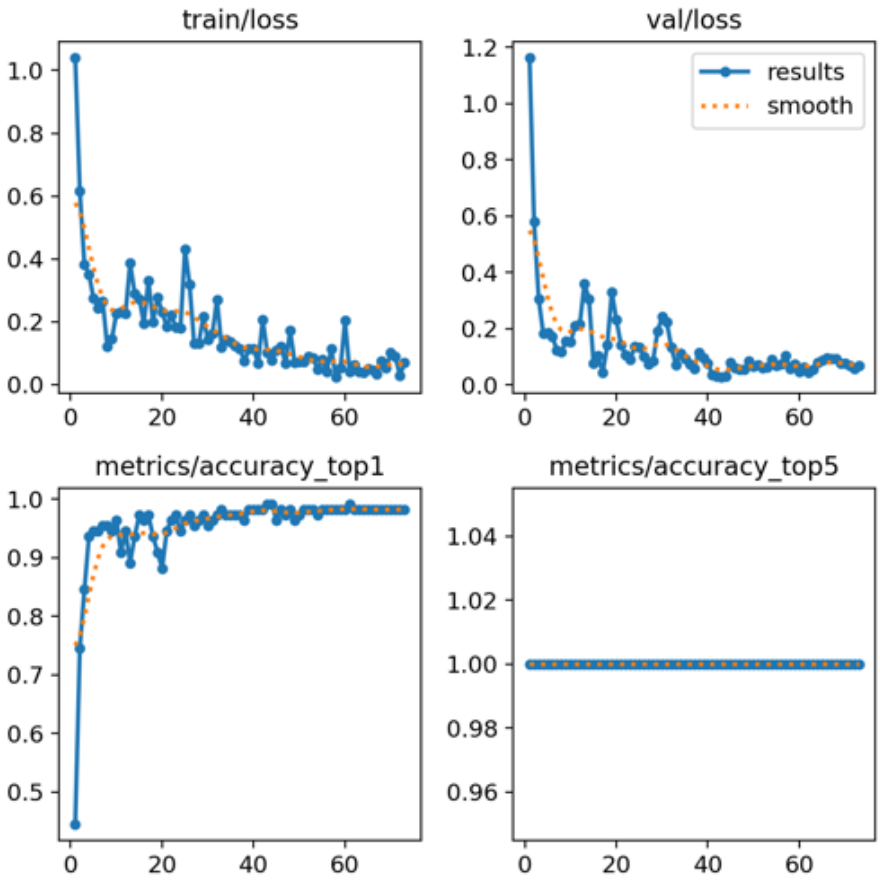

Figure 12. Moisture detection training results.

Rising damp demonstrated perfect classification: all 31 samples were correctly identified, demonstrating the ability of the model to recognize the characteristic salt damage patterns characteristic of capillary moisture movement. The absence of background misclassification and the clear dominance of diagonals in the confusion matrix indicate robust feature learning and effective recognition of different moisture damage mechanisms.

Side-by-side comparison of the ground truth results and the model identification results demonstrate the comprehensive ability of the moisture detection system to detect and classify different types of moisture damage in different building environments. Figure 14 shows the detection results, which show excellent performance in recognizing condensation formations on windows and cold surfaces, accurately identifying the characteristic water droplet formations and surface moisture films that distinguish condensation from other types of moisture. The Penetrating Damp Detection System has proven to be highly effective in accurately identifying water stains, discoloration and signs of damage associated with external water penetration through walls, roofs and building envelope defects. The model has demonstrated particular effectiveness in detecting high damp character-

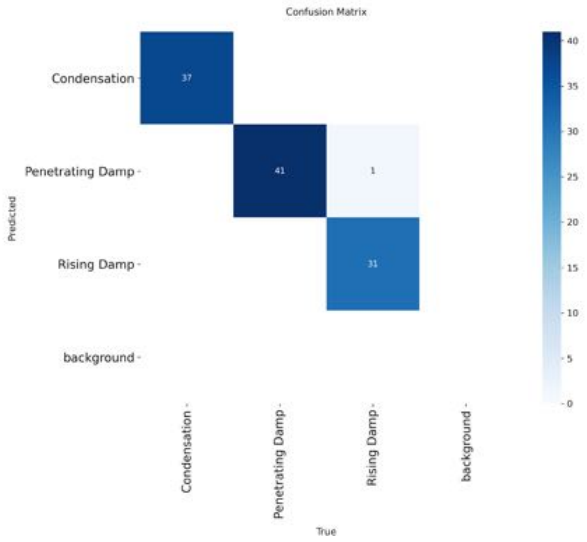

Figure 13. Confusion matrix for moisture classification.

istics, accurately identifying horizontal damp boundaries, salt efflorescence patterns and material degradation associated with capillary movement of moisture upon contact with the ground.

Figure 14. Moisture detection visualization results.

Bounding boxes and classification labels provide precise localization of moisture-affected areas, allowing targeted planning of restoration work and assessment of the extent of damage. High performance enables accurate detection of various surface materials, lighting conditions and moisture levels, from early moisture penetration to advanced damage. The

system effectively differentiated between different moisture sources in complex scenarios where multiple moisture types could coexist in a single structure.

3.2 Mobile application

The defect detection system developed for *Surveyr* successfully demonstrates the real-time crack detection and classification capabilities in a variety of field conditions, validating the practical application of the AI system for on-site building inspection. The interface provides an intuitive mode selection between crack and moisture detection, allowing inspectors to customise the system to suit specific assessment requirements. The crack detection results demonstrate the system's ability to accurately classify structural defects in accordance with BRE Digest 251 standards, with the app accurately identifying an aesthetic concrete floor crack, a serviceable external concrete crack with an appropriate severity rating, and a critical stability-threatening masonry crack in walls. Each classification is accompanied by clear visual feedback and severity indicators, enabling instantaneous decision-making in the field. The mobile version delivers consistent performance across a range of lighting conditions, camera angles and surface textures, demonstrating robust deployment optimisation. The user interface design is focused on fast assessment workflows with one-click detection and immediate display of results, ensuring efficient building inspection work. The mobile application for crack detection can be seen in Figure 15.

The results of the moisture detection function in the mobile application, shown in Figure 16, demonstrate the broad capabilities of the application in detecting and classifying different types of moisture damage in real-life building environments. The interface easily switches between detection modes with clear visual indicators for moisture analysis in line with BRE Digest 245 recommendations. The condensation detection system successfully identified the characteristic signs of moisture on building surfaces, recognising evaporation damage features and thermal bridging effects with the appropriate classification confidence. The high moisture detection system proved to be highly effective in accurately identifying the characteristic horizontal moisture boundaries and material failure patterns consistent with capillary movement of moisture in contact with the ground, as demonstrated by the successful classification of moisture damage to foundations and walls. The penetrating moisture identification system demonstrated high performance in recognising the nature of water penetration through building envelopes, correctly identifying damp spots and structural damage associated with external water penetration around building openings. The mobile app provided consistent detection accuracy across a range of building materials, lighting conditions and humidity levels, from early intrusion to severe damage.

3.3 Web application

The web-based classification platform provides a comprehensive analytical environment for detailed structural assessment and documentation workflows. The clean, professional design of the interface provides a pleasant user experience with intuitive navigation and a clear visual hierarchy, allowing for both quick inspection and detailed analysis. The crack

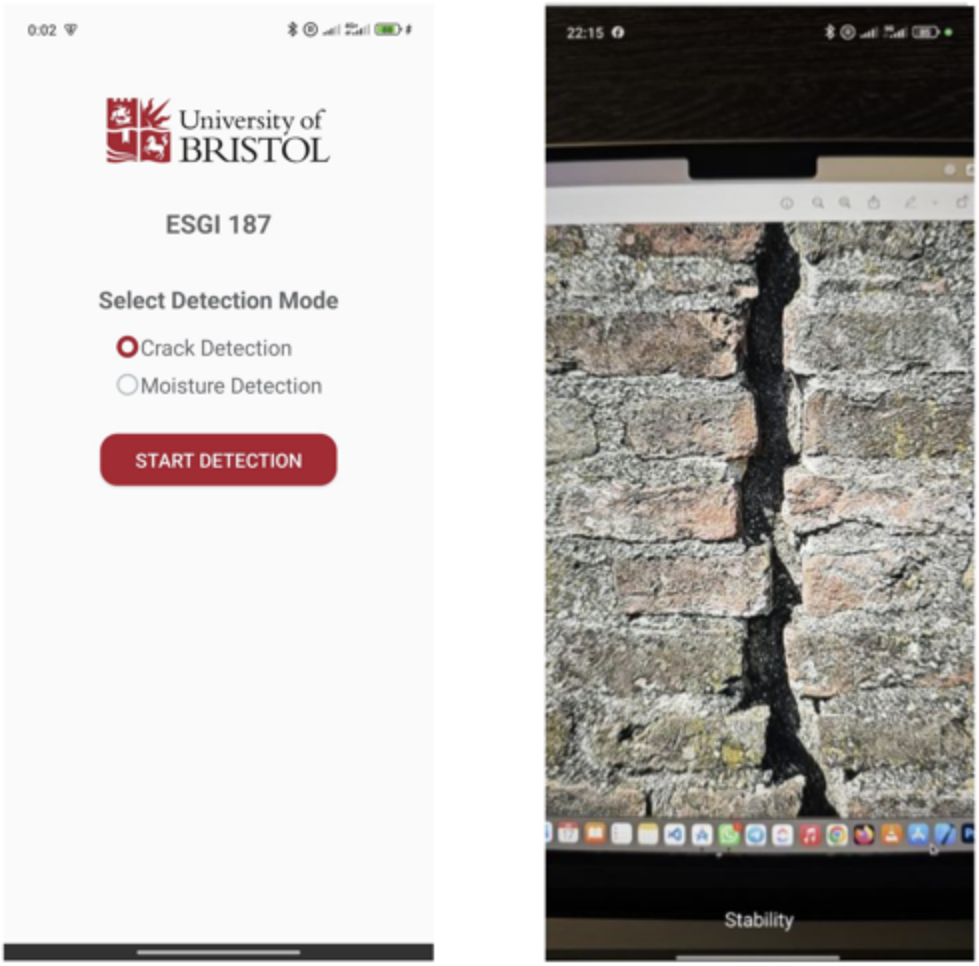

Figure 15. Integration with mobile applications for crack classification.

classification results demonstrate the integration of the system with web technologies, successfully processing uploaded building images and providing detailed damage assessments with a high degree of confidence. An example of the classification of fine cracks around a window frame resulted in 99% confidence in the assignment of an aesthetic category, demonstrating the robustness of the system for professional building inspection. The ability of the platform to process high-resolution images while maintaining real-time response capabilities confirms the successful integration of AI models with the web infrastructure. The results can be seen in Figure 17.

3.4 Roboflow

The Roboflow deployment platform demonstrates successful integration of trained models with a cloud-based inference infrastructure, providing scalable and affordable AI-powered building inspection capabilities. The platform interface provides comprehensive

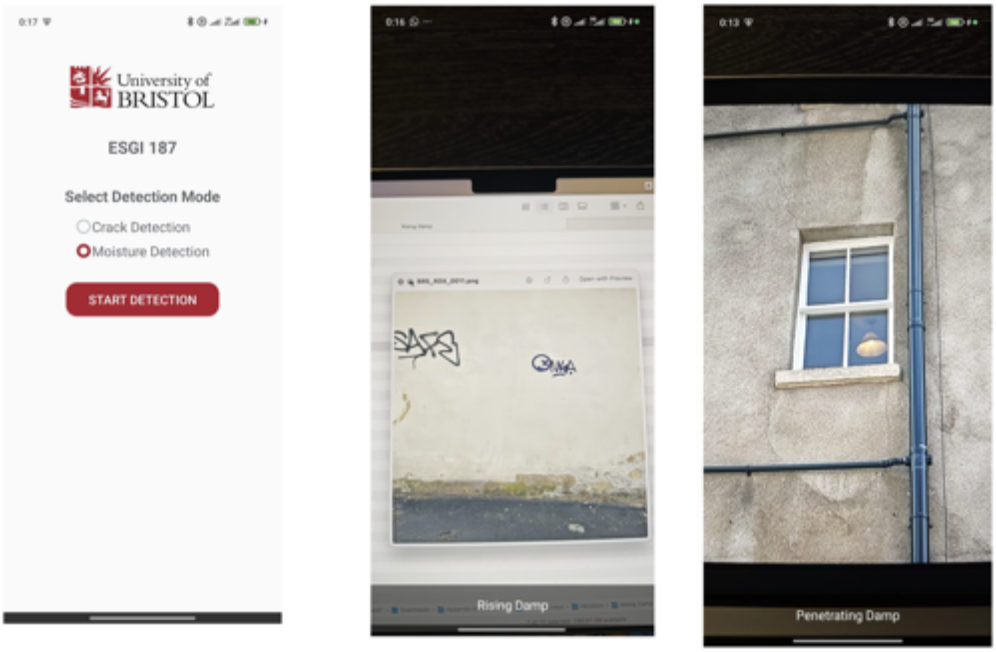

Figure 16. Mobile implementation of the humidity detection system.

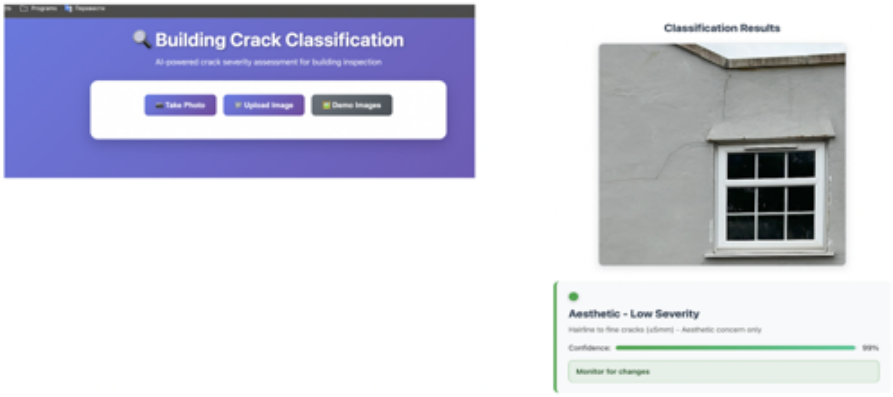

Figure 17. Web platform classification interface.

model management functionality, including test case visualization, confidence threshold customization, and real-time inference capabilities. The humidity detection system example successfully detected increased humidity with 97% confidence, demonstrating the system's high performance in cloud deployment scenarios. The JSON output format provides structured data suitable for integration with existing building management systems, survey databases, and automated reporting workflows. Roboflow results can be seen in Figure 18.

The platform's ability to handle a variety of input methods, including file uploads,

Figure 18. Roboflow model deployment platform.

URL-based processing, and API integration, demonstrates versatility for a variety of deployment scenarios, from individual assessments to large-scale building portfolio analysis. A confidence threshold adjuster allows users to customize the sensitivity based on application-specific requirements, balancing detection accuracy and false positive rates to meet validation protocols. End-to-end prediction results include class identification, confidence scores, and bounding box coordinates, providing comprehensive information for downstream analysis and documentation systems.

4 Conclusion

This research successfully developed and validated a defect detection system developed for *Surveyr*, an integrated AI system for automated building defect detection and classification that addresses critical challenges in building surveying inspection workflows. The multi-architecture approach combining custom CNN, ResNet-50, and YOLO v11 models demonstrated robust performance across crack and moisture detection tasks, achieving high accuracy rates while maintaining compliance with BRE Digest 251 and 245 industry standards. The custom CNN effectively distinguished between cracked and non-cracked surfaces, while ResNet-50 provided reliable severity classification with 72% overall accuracy across aesthetic, serviceability, and stability categories. YOLO v11 delivered exceptional performance with 90% top-1 accuracy and perfect top-5 accuracy, successfully integrating detection and classification in a unified framework. The moisture detection system achieved near-perfect classification across condensation, penetrating damp, and rising damp categories, with confusion matrix results showing minimal inter-class confusion. Practical deployment through mobile and web applications validated the system's real-world applicability, enabling real-time field assessments with intuitive user interfaces and immediate classification feedback.

The successful integration with cloud-based platforms demonstrates scalability for enterprise-level building inspection workflows. While challenges remain in handling severely

imbalanced datasets, particularly for critical stability-threatening defects, the system represents a significant advancement in automated building inspection technology. Future enhancements incorporating Large Language Models for automated reporting, stereo camera integration for precise crack width measurement, and expanded training datasets will further advance the practical utility of AI-powered structural assessment systems in professional building inspection practice.

Surveyr Part 2: Cracks!!!–Unsupervised.

I. Starikova[1] and M. Zyskin[2]

[1]Slovak Academy of Sciences
[2]Consultant
maximzyskin@gmail.com

Contents

2 Crack detection via unsupervised image analysis

Part 1 of the report is focusing on supervised learning methods, requiring collection and labeling of training sets for training image detection neural network, subsequently used for Yolo-like segmentation search, scoring bounding box jointly with the detection (thus enabling to recognize learned objects in a larger image full with competing details). Some public neural networks detecting cracks are already trained and made available, for example on the Roboflow platform. While Yolo method is published as a research paper, and also can be these days called from Python as an open source library (we have done so as part of the project) — and can be coupled in the same Python code via API with a public trained detection model deployed somewhere. e.g. on Roboflow (we have tried this, though our Part 2 team have not pursue it to the end), creating a possible route to solve the Surveyr problem. Further details and discussion of this line of approach have been explored and presented in the spectacular Part 1 of the report by the Kazakh team.

However , this is really not our point here. In this part 2 of the report, we would like to explore primarily image analysis and (unsupervised) learning methods, in as much as they apply to crack detection, and in as much as they seem promising for developing image analysis and machine learning methods.

2.1 Stand-alone cracks

'Old school' methods of edge detection involve only simple operations on pixilated image, essentially based on discrete version of a derivative, or convolution with a Gaussian kernel and differentiating the result, and then studying regions where the norm of the gradient is large , as well as looking at second derivatives, such as sign of Laplacian, to distinguish narrow black feature from narrow white feature; eigenvalues of the Hessian, to verify that that has one large eigenvalue and one small eigenvalue, the latter corresponding to a direction where gray level does not change much and orthogonal to direction where it changes a lot.

Such edge detection methods would be very effective if a well-isolated crack is seen, for example:

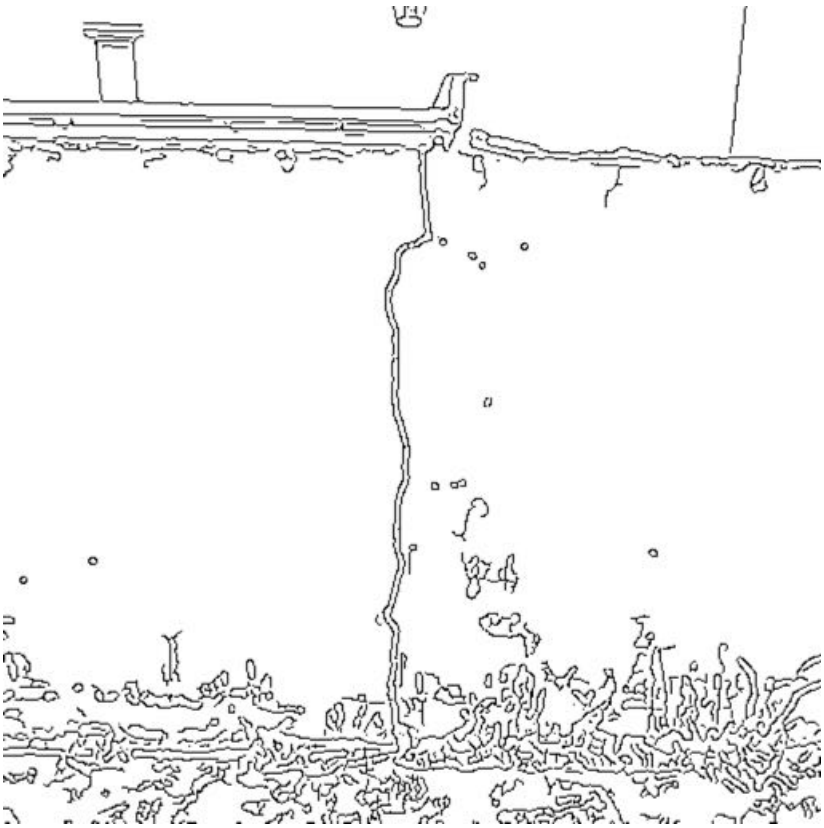

Figure 19. Edge detection applied to gray scale image. Vertical long crack is clearly seen. Building features (top), grass and wall edge (bottom) are also seen

Well-isolated cracks are also picked up by some public trained and deployed models, see for example Fig. 20.

Another public model, Fig. 21, with name ending in ../crack-detection-ypnwo/v3, was also successful, however we have not saved, and could not now find, the link to it now:

The other few models we have also tried, some of them with larger training set, have failed to detect crack in this image. We observed that although image preparation like edge detection helps a lot for a human to see cracks, this does not necessarily extends to models pre-trained on full untreated images. We expect though that training on prepared images could be more successful, as in a way there is less to learn.

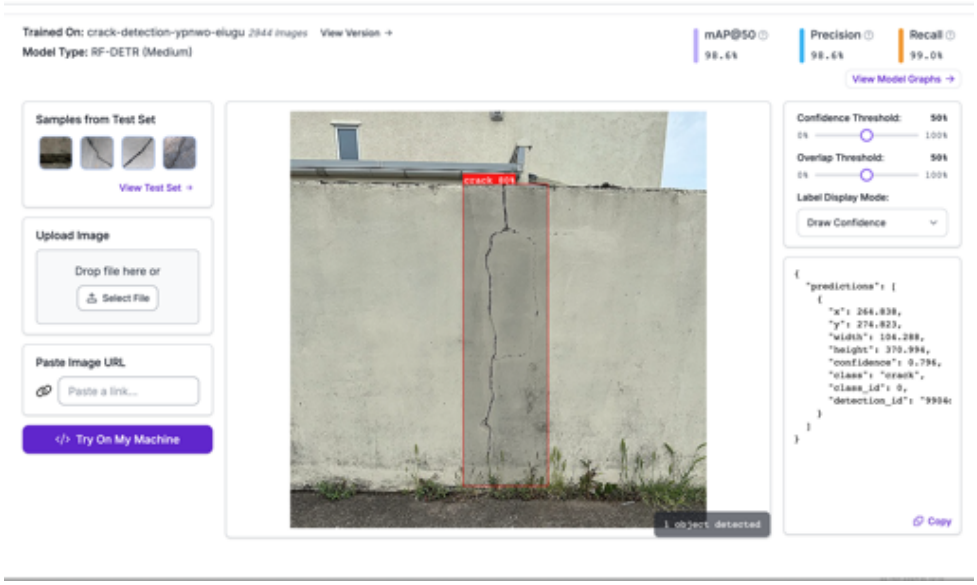

Figure 20. Public trained model deployed on Roboflow finds and identifies the feature as crack, correctly. The model link: https://universe.roboflow.com/roads-fgydm/crack-detection-ypnwo-eiugu/model/1, checked on Sept 1, 2025.

Another public model with name ending in ../crack-detection-ypnwo/v3, was also successful, Fig. 21 (when tested on Aug. 29, 2025), however we have not saved, and could not find the full link to it, at this time.

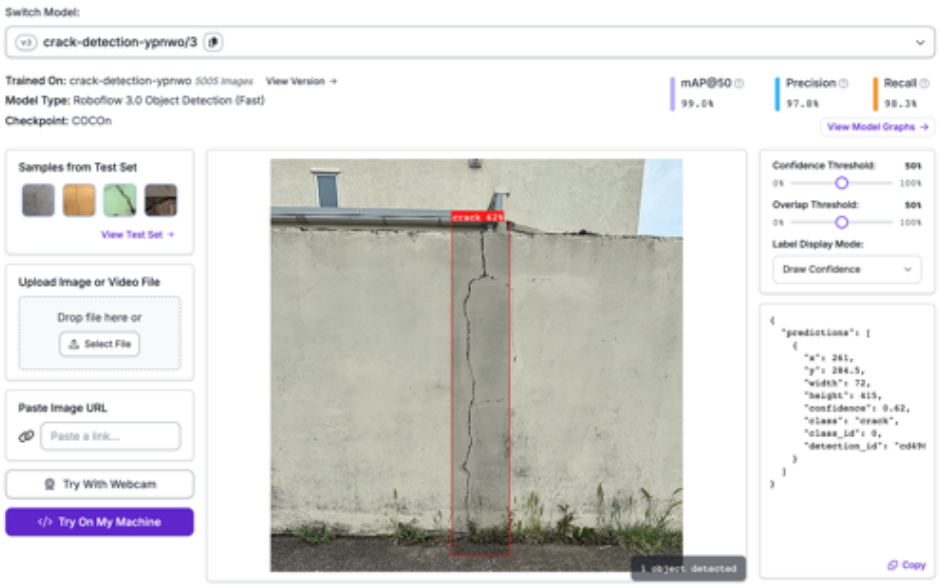

Figure 21

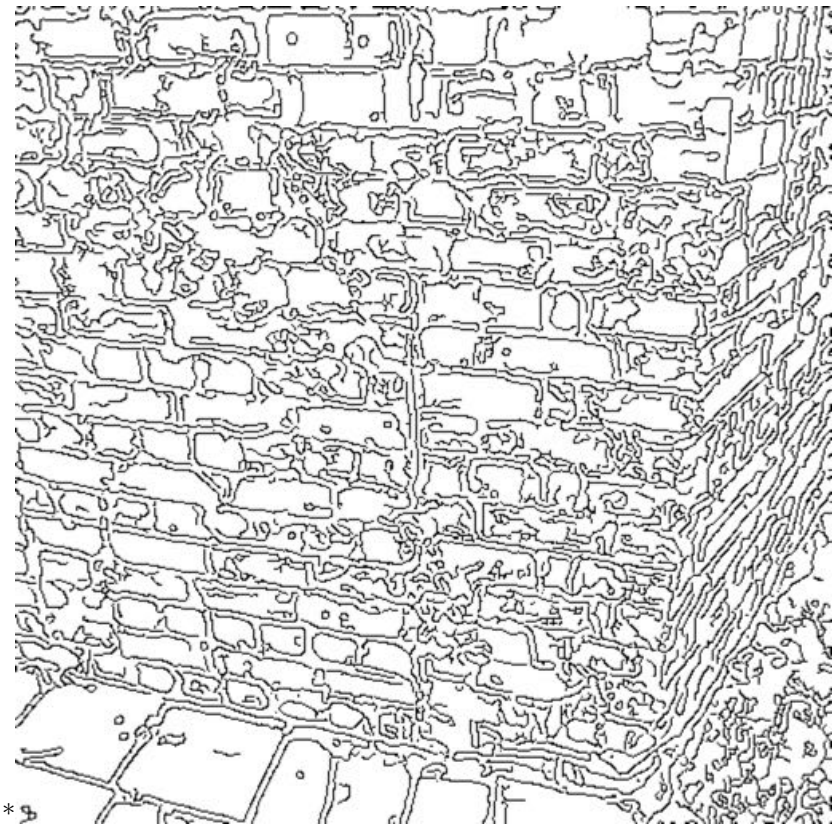

Figure 22. Edge detection on a brick wall seems a bit loaded...Can you see the crack?

2.2 Bricks !!!

2.3 Pre-processing

Cracks through mortar in brick walls provides more challenge. Edge detection will pick up not just the crack, but also bricks/mortar boundaries, as well as other sharp features, and may look as so: We will outline below an unsupervised detection procedure, which would not require extensive collection and labeling of training sets (or alternatively can assist in supervised learning, by simplifying and reducing amount of non-essential features in the imagery. Also, such reduced imagery can be generated synthetically by a computer, with actual raw images not necessarily required. Our approach consists of 2 main steps, broadly described as (i) a filtered edge detection (ii) curves fitting and post-selection.

We have examined several filtered detection approaches. 1. Converted to grayscale image convolution with a Gaussian kernel, averaging over an area on a selected scale. With a good scale choice, this softens/washes out mortar edges, relative to cracks, Fig. 23, right. 2. Follow-up adjusting of the image gray scale to its full 0 to 1 range is suppressing lighter brick and mortar region relative to the crack, Fig. 23, (a). Adjusting contrast also have a similar effect. 3. Further, it proved very effective to suppress lighter gray areas altogether, as the crack is the darkest, as it is not reflecting visual light back.

Keeping only dark areas (we kept values below 0.05), followed by attempted removal of small features (we have not observed a significant effect) is shown in Fig. 24, left. Frangi vesselness filter, based on eigenvalues of the matrix Hessian H, and developed for blood vessel detection, is applied to the result of the previous step.

For a 2D image, the Hessian matrix H at point (x, y) is:

$$H = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial x \partial y} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} \tag{2.1}$$

Let λ_1 and λ_2 be the eigenvalues of H with $|\lambda_1| \leq |\lambda_2|$. The "blobness" parameter is defined as $R_B = \frac{|\lambda_1|}{|\lambda_2|}$, and the "structureness" parameter as $S = \sqrt{\lambda_1^2 + \lambda_2^2}$. The filtered result, in the version designed to spot dark cracks (low grayness score) on light background (higher grayness score), is then given by

$$V(s) = \begin{cases} 0 & \text{if } \lambda_2 < 0 \\ \exp\left(-\frac{R_B^2}{2\beta^2}\right)\left[1 - \exp\left(-\frac{S^2}{2c^2}\right)\right] & \text{otherwise} \end{cases} \tag{2.2}$$

In a multiscale implementation, a maximum over scales may be also taken,

$$V_{\text{MS}} = \max_{s_{\min} \leq s \leq s_{\max}} V(s). \tag{2.3}$$

In Fig. 24 (b) we show the result of applying Frangi filter to the raw image converted to gray scale (and slightly smoothened with a Gaussian 3 pixels-wide kernel). In Fig. 24 (c) and (d) we show results of edge detection and Frangi filter respectively post-applied to Fig. 24 (c) which keeps only darkest pixels. We observe that after any of the 4 preprocessed methods in Fig. 24, location of the crack can be seen clearly with a human eye, to a various degree of visual contrast.

Figure 23. Left, convolution with a 2 pixels wide Gaussian kernel is washing out mortar edges. Right, follow-up gray scale adjustment to the full 0 to 1 range.

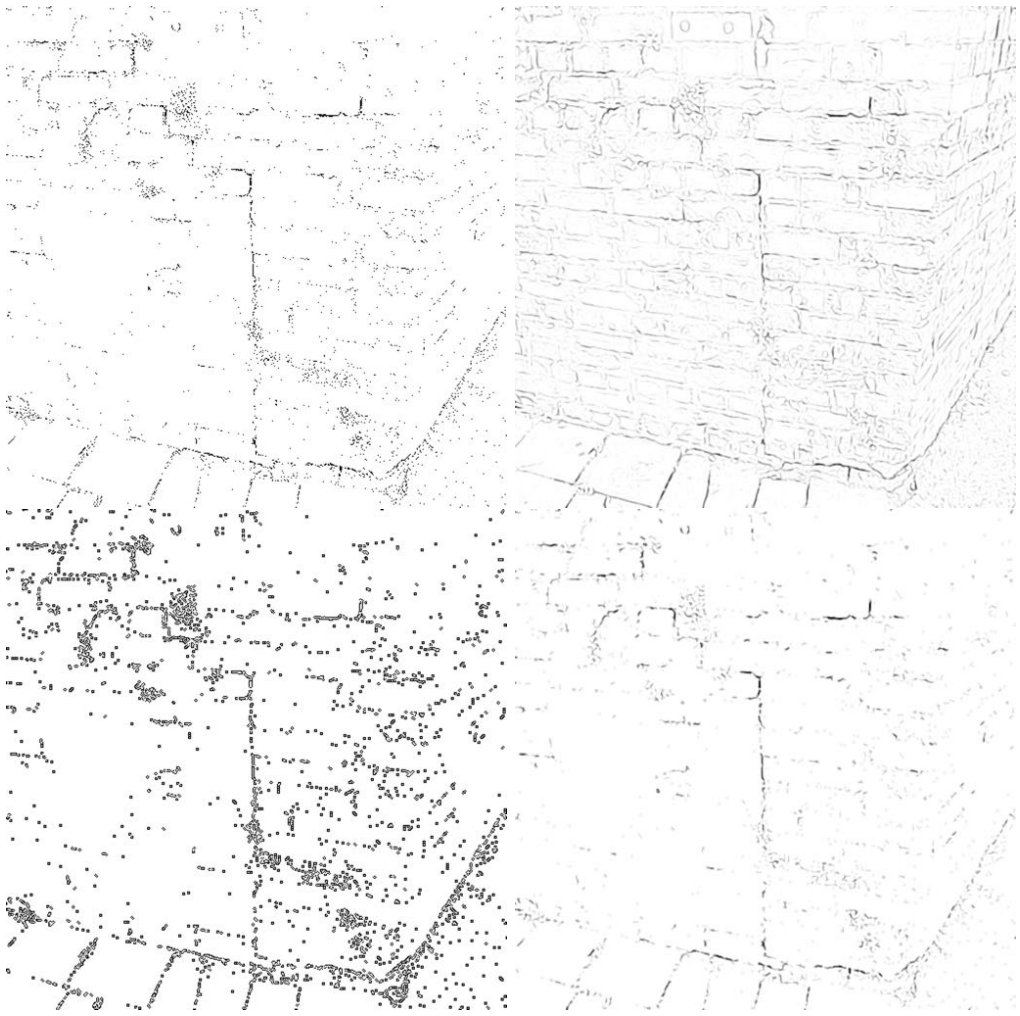

Figure 24. (a) Top left, keeping only dark, gray score ≤ 0.05 regions, followed by small features removal. (b) Top right, Frangi filter with $\beta = 0.5$, $c = 0.1$ applied to the raw image converted to gray and smoothed by 3 pixel Gaussian filter (c) Bottom left, edge detection applied to (a); (d) Bottom right, Frangi filter applied to (a)

2.4 Contours detection

We would like now for the computer to highlight a guess on crack location, by surrounding it by a contour, say. We found that in the preprocessed image the crack pixels are not connected enough for this to work well, so we first dilated dark regions, transforming Fig. 25, left to a thickened version, right. Dilation is done as the Minkowski sum of the negated image with a disk of 4 pixels radius (see the Appendix).

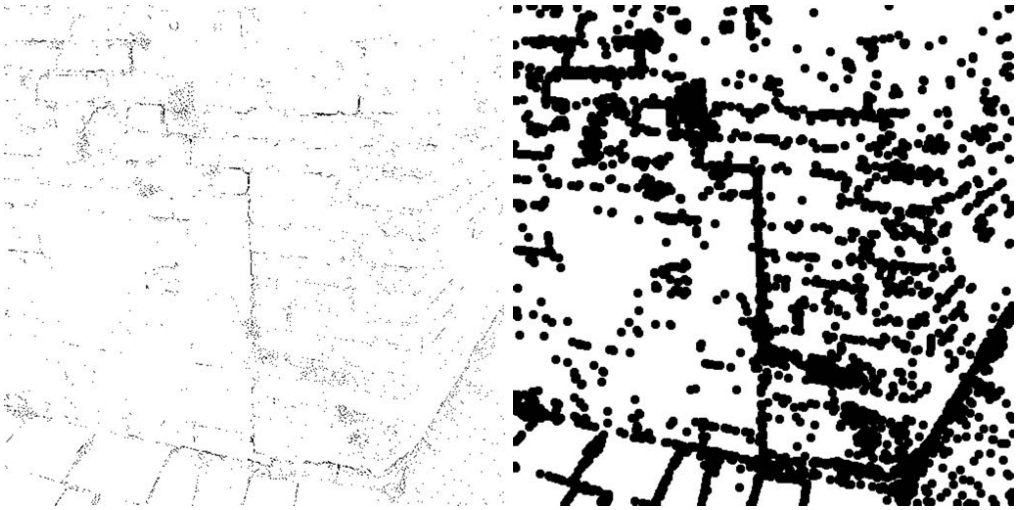

Figure 25.

This is followed by the resulting image binarized into zero and one (black and white, if not like this already, and a contour tracing algorithm is applied to trace the boundary of those by closed polygonal curves (which may involve joining segments, or cutting to avoid self-intersections) Selection is then made among such contours, which may be based on arc length, or largest distance among pixels enclosed. Top contours found by Mathematica, ordered by arc length, are shown in Fig. 26, left. Top contours found using open source vision library and Python are shown in Fig. 26, right.

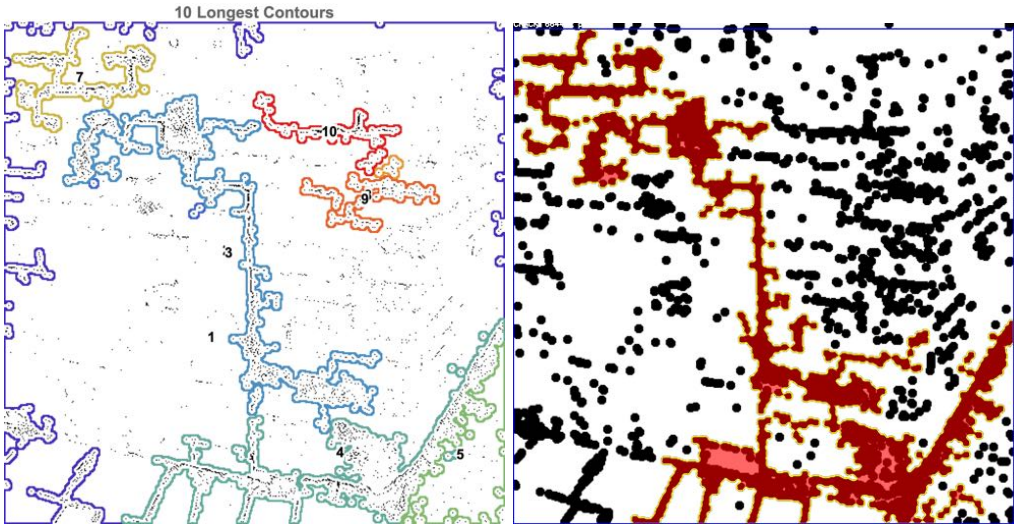

Figure 26. Left: top-length contours fitted on thickened image ;right: final result of an open source Python code contour search; see the Appendix for details.

Further development of the contour tracing method can distinguish algorithmically between contours bounding cracks through bricks mortar (staircase-like) from contours bounding bricks (quadrilateral) , or mortar still present despite pre- processing (lattice-like). We have not tried to do so, in this project of a short duration and no budget, applying already available contour tracing methods instead.

3 Appendix

3.1 Open Source Vision Library

A version of the Python code searching for highest-scoring contours resulted in the finding shown in Fig. 26.

Only free tools have been used in the code (Open Source Vision library cv2 and the standard Python libraries).

Figures below are the intermediate steps produced by the code, and the final result.

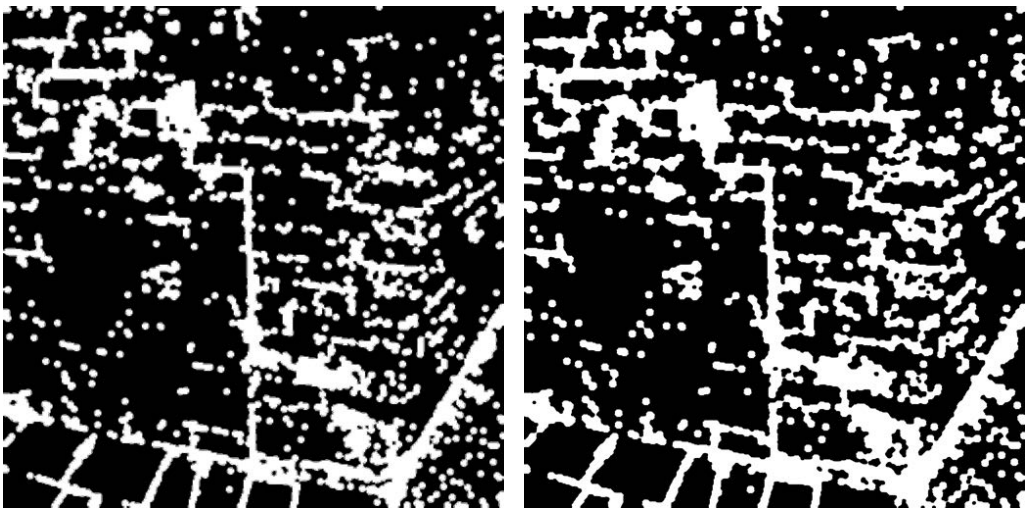

Figure 27. 1. Darkest Regions-left; 2. Connected Regions -right

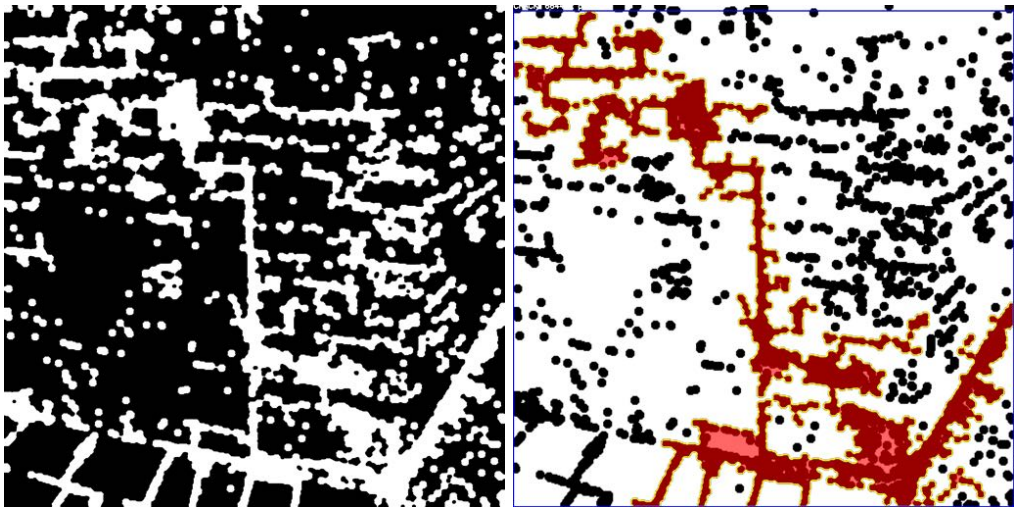

Figure 28. 3. Largest Components-left; final result-right.

3.2 Dilation, by a Minskowski sum with a disk

Let the binary image be represented as a set of foreground pixel coordinates:

$$I = \{\mathbf{p} \in \mathbb{Z}^2 | \text{Pixel at } \mathbf{p} \text{ is foreground}\}$$

The structuring element `DiskMatrix[4]` is a discrete disk of radius 4, defined as the set:

$$B = \{\mathbf{x} \in \mathbb{Z}^2 : \|\mathbf{x}\| \leq 4\}$$

This creates a 9×9 matrix (diameter $= 2 \times 4 + 1 = 9$), where the origin is at the center.

Dilation Operation

The dilation of image I by structuring element B is given by the Minkowski sum:

$$I \oplus B = \{\mathbf{p} + \mathbf{b} | \mathbf{p} \in I, \mathbf{b} \in B\}$$

Equivalently, using the set-theoretic definition for a binary image:

$$(I \oplus B)(\mathbf{s}) = \bigvee_{\mathbf{b} \in B} I(\mathbf{s} - \mathbf{b})$$

where $\bigvee$ denotes the logical supremum (OR) operation.

The procedure was applied to the negated image, negated back after the procedure, so that it was the dark areas which was dilated.

References

[1] Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 32(5), 361-378.

[2] Flah, M., Nunez, I., Ben Chaabene, W., & Nehdi, M. L. (2021). Machine learning algorithms in civil structural health monitoring: a systematic review. *Archives of Computational Methods in Engineering*, 28(4), 2621-2643.

[3] Dorafshan, S., Thomas, R. J., & Maguire, M. (2018). Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Construction and Building Materials*, 186, 1031-1045.

[4] Nature Scientific Reports. (2025). Comparative analysis of deep learning models for crack detection in buildings. DOI: 10.1038/s41598-025-85983-3.

[5] Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2024). Road crack detection using deep convolutional neural network. *Computer Vision and Image Understanding*, 242, 103945.

[6] Nature Scientific Reports. (2024). Deep neural networks for crack detection inside structures. DOI: 10.1038/s41598-024-54494-y.

[7] Dung, C. V., & Anh, L. D. (2019). Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 99, 52-58.

[8] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 779-788.

[9] Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics. Available: https://github.com/ultralytics/ultralytics

[10] Li, H., Xu, H., Tian, X., Wang, Y., Cai, H., Cui, K., & Chen, X. (2023). Bridge crack detection based on SSC-YOLO with multi-scale feature fusion. *Measurement*, 213, 112707.

[11] Balaras, C. A., & Argiriou, A. A. (2002). Infrared thermography for building diagnostics. *Energy and Buildings*, 34(2), 171-183.

[12] Edis, E., Flores-Colen, I., & de Brito, J. (2014). Passive thermographic detection of moisture problems in façades with adhered ceramic cladding. *Construction and Building Materials*, 51, 187-197.

[13] Garrido, I., Lagüela, S., Arias, P., & Balado, J. (2018). Thermal-based analysis for the automatic detection and characterization of thermal bridges in buildings. *Energy and Buildings*, 158, 1358-1367.

[14] Spencer Jr, B. F., Hoskere, V., & Narazaki, Y. (2019). Advances in computer vision-based civil infrastructure inspection and monitoring. *Engineering*, 5(2), 199-222.

[15] Xu, H., Su, X., Wang, Y., Cai, H., Cui, K., & Chen, X. (2019). Automatic bridge crack detection using a convolutional neural network. *Applied Sciences*, 9(14), 2867.

[16] Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road damage detection and classification using deep neural networks with smartphone images. *Computer-Aided Civil and Infrastructure Engineering*, 33(12), 1127-1141.

[17] Han, S., Mao, H., & Dally, W. J. (2015). Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. arXiv preprint arXiv:1510.00149.

[18] Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531.

[19] Özgenel, Çağlar Fırat. (2019). Concrete Crack Images for Classification. Mendeley Data, V2. DOI: 10.17632/5y9wdsg2zt.2

[20] Building Research Establishment. (2018). *BRE Digest 251: Assessment of damage in low-rise buildings*. BRE Press.

[21] Building Research Establishment. (2016). *BRE Digest 245: Rising damp in walls: diagnosis and treatment*. BRE Press.

[22] https://www.kaggle.com/datasets/arunrk7/surface-crack-detection