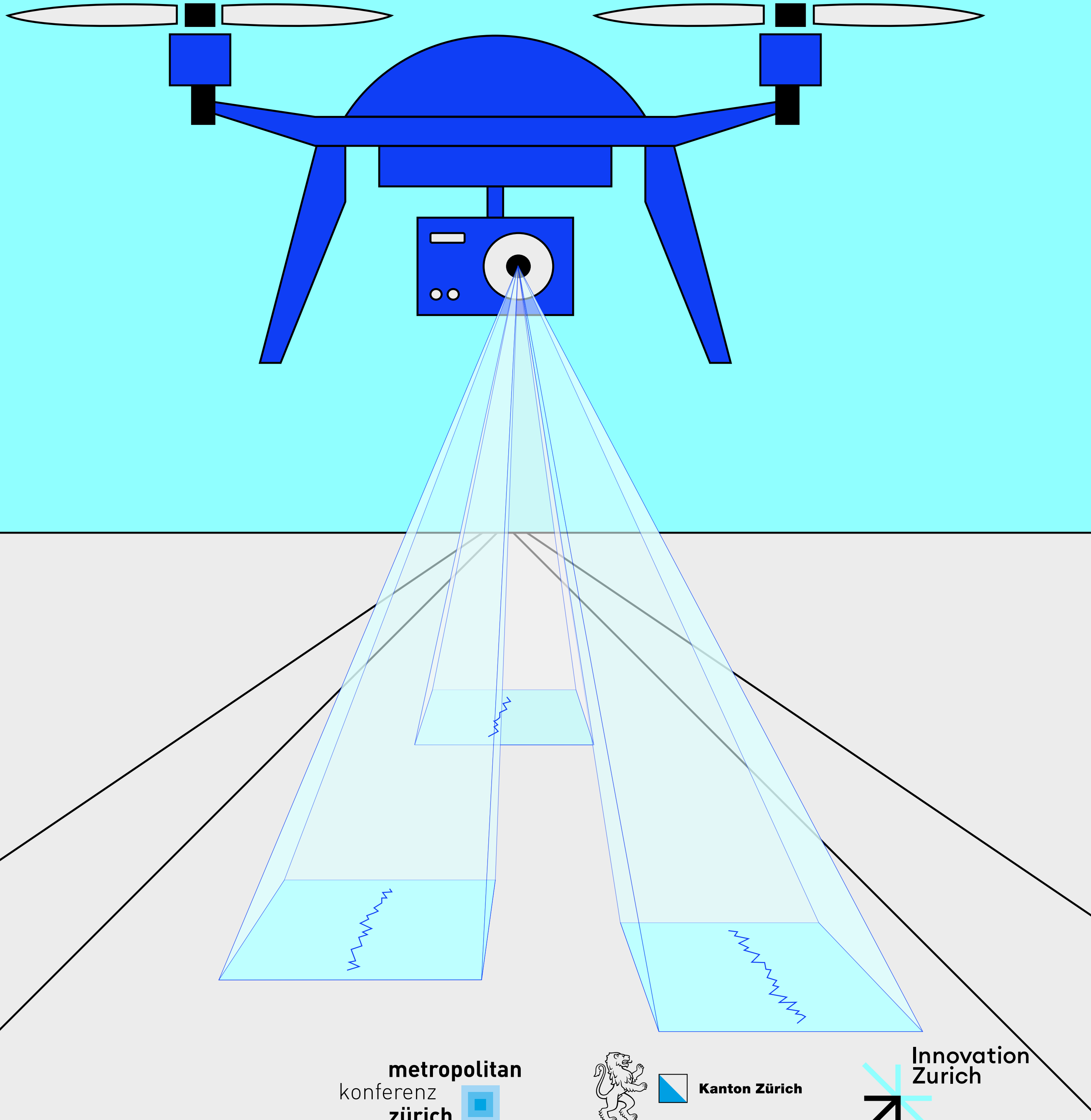
Innovation Sandbox for Artificial Intelligence

Automated Infrastructure Maintenance

Drone Inspections with Computer Vision



metropolitan konferenz zürich

Kanton Zürich

Innovation Zurich

Infrastructure maintenance of roads, bridges and dams offers great potential for use of artificial intelligence (AI). AI-based image recognition can systematically and automatically detect the tiniest of cracks or defects. Infrastructure operators still largely carry out inspections manually. Within the scope of the Innovation Sandbox for Artificial Intelligence (AI), IBM Research and pixmap gmbh implemented a pilot project on the Dubendorf Air Base, to assess the potential of AI-based inspections. In the project, a drone created high-quality imagery of the runway using AI to automatically detect any defects. The project findings are being employed to advance the use of AI for inspection and maintenance of further infrastructure elements. The imagery is being made available to other innovation stakeholders. By way of this project, the collaboration between public administration, military, research and private industry is contributing to the further development of the Zurich Metropolitan Area as an international hub for AI.

Innovation Sandbox for Artificial Intelligence (AI)

This document was created within the scope of the Innovation Sandbox for Artificial Intelligence (AI). The sandbox is a test environment for the implementation of AI projects from various sectors. This broad-based initiative involving public administration, industry and research, is designed to promote responsible innovation by allowing the project team and participating organisations to collaborate closely on regulatory questions and enabling the use of novel data sources.
More Information

Table of Contents

I.

Challenges of manual inspections

The maintenance and inspection of infrastructure elements, such as roads, bridges and dams, is of immense significance for public safety and for maintaining economic activity. That being said, infrastructure operators are faced with a massive and complex task: Switzerland's road network is more than 84,000 km long.[1] More than 40,000 bridges span gorges, valleys and rivers.[2] With over 200 major dams and thousands of small-scale ones, Switzerland is the country with the highest density of dams in the world.[3] This raises the question as to how such a vast quantity of infrastructure elements can be monitored efficiently. Numerous inspection processes are based on manual surveys of cracks, defects and other irregularities in infrastructure. This leads to a series of challenges ranging from lack of efficiency to human error. The potential offered by AI automation has not yet been tapped. In chapter one of this document, the four core challenges of manual inspections will be explored, followed by an elucidation of the potential of automated inspection through image recognition.

Lack of efficiency

Manual inspections require that a person physically inspect the entire surface of an infrastructure, e.g. runways or road sections. As a result, manual inspections not only consume a great deal of time and human resources, but also tend to make maintenance cycles longer than they should be. The process of inspecting infrastructure elements by physically walking or driving across them is the cause of delays which could be avoided with the help of automated systems. Manual inspections also tend to be very costly.

Poor documentation

Whereas manual inspections do provide data on defects and issues, they often lag behind what is technically possible today. In many cases, instead of a comprehensive digital image (**digital twin***) that continually documents the condition of an infrastructure over a longer period, a list of problematic areas or defects is compiled. The location of the defects is often only roughly indicated, making it harder to fix the problem. The absence of a digital model with automated recognition also means that data is often not integrated into operating systems and thus cannot serve as a basis for decision-making.

Human error

Humans make mistakes which can be due to concentration problems, especially in the context of repetitive and tedious tasks, such as inspecting long stretches of roads. Furthermore, there is the issue of consistency: what may qualify as a defect to one person may be deemed insignificant by another. Unrecorded defects or mix-ups can lead to expensive repairs or even infrastructure risks.

Safety hazard

Manual inspections often harbour risks for inspection staff, especially when tasks involve working in exposed locations, such as in the vicinity of

* The highlighted terms in the text are explained in the glossary.

high-voltage lines or in high places. Accidents occur every year as a result of manual inspections. Whereas safety equipment and training can help to reduce risks, the question remains as to whether it makes sense to expose people to such risks, seeing that technological alternatives are available.

Potential for automated inspections

The mentioned challenges point to the major potential of automated inspections. The progress made in image recognition technology is paving the way to new opportunities to overcome some of these challenges and to significantly increase efficiency and precision of inspection processes. In this project, the stakeholders from public administration, research and industry joined forces to contribute to automated infrastructure maintenance. IBM Research submitted a project proposal to the Innovation Sandbox for AI in spring 2022. The idea was to make use of the experience gained from previous projects with AI-based bridge inspections and to expand this knowledge to airport runways. The Innovation Sandbox for AI supported this innovation initiative through providing high-quality image data and by engaging the Dubendorf Air Base as a project partner from the innovation ecosystem of the Zurich Metropolitan Area. The creation of imagery was done in collaboration with pixmap gmbh, a company specialised in conducting inspections and surveys with drones and flying robots.

This report is divided into the following sections: chapter two describes the process and insights gained from the drone missions carried out by pixmap gmbh. Chapter three gives an overview of the automated evaluation of the imagery using IBM Research image recognition technology. A conclusion is drawn in chapter four and several potential areas of action shown as to how to advance automated inspections of infrastructure elements in the Zurich Metropolitan Area. A description of the technical details of image recognition by IBM Research is provided in the appendix.

[1] Infrastructure and network length| Federal Statistical Office (admin.ch)

[2] Across rivers and gorges – Explora (ethz.ch)

[3] Swiss dams – second to none | House of Switzerland

II.

Use of drones to create imagery

Comprehensive imagery is the bedrock for automated detection of cracks and defects in infrastructure elements. There are various methods and sources from where such data can be drawn. In some cases, use of freely accessible satellite images will prove sufficient, e.g. when damage to infrastructure is clearly visible at low resolution. However, if high-resolution imagery is needed for an in-depth inspection, there is no way around a more targeted image capturing. Images can be captured from a ground vehicle or from the air, e.g. with a drone. In the case at hand, pixmap gmbh used drones. Drones have the considerable advantage of being able to produce systematic images that are accurate down to the last centimetre, and that are repeatable and in high resolution. Furthermore, drones are easy to transport and can be operated at relatively low cost.

The Dubendorf Air Base put a more than 2.8 km long runway at the disposal of the pilot project. The airport operator defined a representative runway section of 200x40m. pixmap gmbh took pictures of the area in the required maximum quality by means of a drone, and subsequently made the data available to IBM Research for analysing. The very high resolution made it possible for IBM's AI team to evaluate whether less high-resolution imagery would have been sufficient for an automated analysis of the cracks and defects. This is especially relevant in view of the operational use of automated inspections, e.g. if optimum quality is too costly or if data collection were to prove too time-consuming for regular maintenance work.

In the given context, planning the use of drones meant taking several challenges into account.

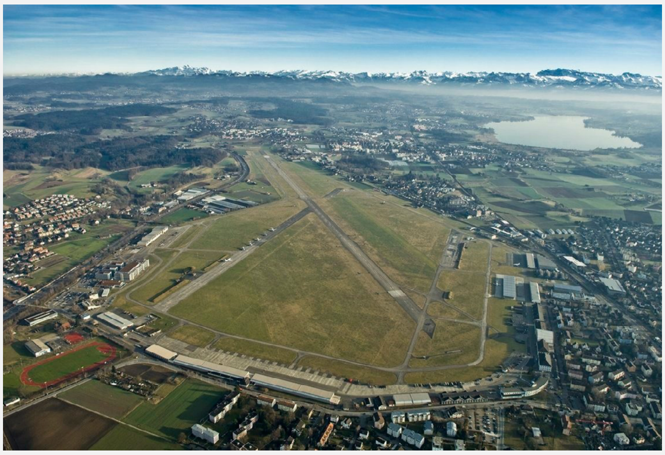


Figure 1: 2.8 km long runway of the Dubendorf Air Base in the Canton of Zurich

Regulations: depending on the mission and location, there are regulatory requirements applicable to the use of drones. With the adoption of the drone regulations of EU/EASA on 1 January 2023, statutory provisions are even stricter now. Drone operators must substantiate with an operating licence that risks for third parties on the ground (Ground Risks) as well as collisions with other aircraft (Air Risks) can, with all likelihood, be avoided. In the present case, the runway was closed and the flight altitude of the drone was limited to maximum 10 m, so that other airfield operations with helicopters were not affected.

Drone/camera requirements: a high-quality camera on a drone with precision **GPS** is required for sub-millimetre resolution images. Modern drones can be equipped with full-frame cameras (36x24 mm sensor size) and resolutions of 40 to 100 MP. It is also important to have a camera with short shutter speeds so as to avoid motion blurs, and to have a short trigger interval, in the present case 0.7s. To be able to steer the drone with accuracy down to the centimetre across the runway, it has to be equipped with a special GPS system (**RTK-GNSS**) and, furthermore, be able to fly autonomously according to pre-defined waypoints.

Dependency on weather: note should be taken that these sorts of missions are weather-dependent. The runway must be dry, the drone must not cast any shadow on the captured surface, and strong gusts of wind need to be avoided. Therefore, reserve dates/times are important when planning drone footage.

In the project at hand, pixmap gmbh was able to implement the drone missions as planned on 13 May 2023.

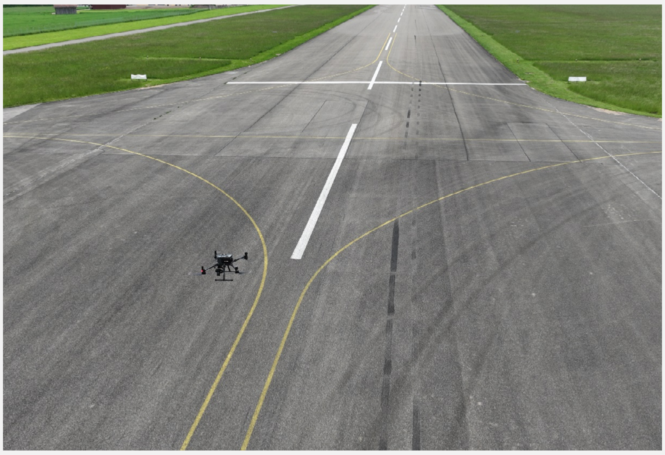


Figure 2: **Quadcopter** flying above the runway at ca. 5 m altitude

For the sake of achieving a data basis that is as broad as possible, the main mission was slightly extended i.e. three different missions with varying collection parameters were flown.

Mission 1: a new resolution level

The typical resolution of drone images in the survey area is between 1 cm and 3 cm. However, this particular mission went far beyond that. The required maximum resolution of 0.25 mm vis-à-vis 1 cm corresponds to a factor of 40 or 1,600 times more pixel points per unit area. This led to extreme flight parameters: equipped with a high-performance camera, the drone had to operate at a flight altitude of just 3 m. A very slow flight speed of 0.7 m/s and a picture taken every 0.7 s were needed to record slightly overlapping images and to avoid motion blur. This resulted in very long flight times totalling 2 hours with approximately 11,500 pictures taken.

Mission 2: lower resolution as an alternative

Whereas the focus of mission 1 (M1) was on maximum resolution, mission two (M2) centred on scalability, hence, on the option to later be able to scan and capture an entire runway. A slightly lower resolution of 0.75 mm allowed for a ten-fold faster capturing time, with a correspondingly smaller data volume.

Mission 3: focus on mapping

The aim of this mission (M3) was to create an even more comprehensive overall image of the runway using photogrammetry. For this purpose, all of the individual images were converted into one, georeferenced **orthophoto** and a digital elevation model using a special software (Pix4Dmapper). This is how, in lieu of thousands of individual pictures, a single, undistorted overall image was achieved that could then be further analysed. Photogrammetric capturing is, however, only possible with significantly higher overlaps of the individual images, which, in turn, reduces the resolution achievable. If the aim is to take photogrammetric images of the entire runway, a resolution of approx. 1.5 to 2 mm can be achieved with today's technology. For the runway section defined in this project, a resolution of 0.6 mm was achieved.

Conclusions drawn from the drone missions

The results are positive - all of the missions were completed successfully on first attempt and allowed for high-quality data to be generated straightaway, without any gaps in coverage. The comprehensive analysis conducted by IBM Research (see chapter three) shows that the aim of automatically detecting runway defects is achievable even with a slightly reduced resolution of 0.75 mm (M2). This means that the method applied here can already be used today for entire runways or longer road sections. Such missions will need to be planned carefully with the necessary know-how, and carried out with high-quality equipment focussing on goals that have been accurately defined by the operator. As drone technology continues to move forward, more advanced requirements such as photogrammetric capturing in the sub-millimetre range will soon be possible, with effort and expenditure tending to decrease. Use of drones to inspect infrastructure elements is undoubtedly of major significance.

	Mission 1 (M1)	**Mission 2 (M2)**	**Mission 3 (M3)**
Objective	Best resolution	Scalability to entire runway	Mapping/ photogrammetry
Resolution	0.25 mm	0.75 mm	0.60 mm
Flight speed	0.7 m/s	4.2 m/s	1.1 m/s
Flight time	120 min	10 min	60 min for ca. 1/3
Number of images taken	~ 11,5000	~ 1,200	1 **orthophoto**

Table 1: Comparison of the three drone missions

III.

Automated visual inspection with AI foundation models

IBM Research used high-resolution imagery from pixmap's three drone missions to develop automated AI-based inspection methods.

Advances in **deep learning** are enabling more and more applications in image recognition that were previously considered impossible. Today's fast-paced technological developments are driven by the availability of **annotated data** and specialised hardware such as **GPU** (graphics processing units), which facilitate the training of AI models. This method falls into the category of machine learning and learns from large amounts of data. For this project, the latest advances in deep learning were used to develop technology that detects small cracks in the infrastructure.

One of the main challenges of the project was that annotated data on civil infrastructure is rarely publicly available and most image data does not reveal any visible defects. Solutions like **few-shot learning, transfer learning** and **self-supervised learning** are being explored to overcome this challenge.

IBM Research Zurich brought specialist expertise in the field of automated visual inspection of civil infrastructure to the project, particularly in the domain of concrete bridge pillars. The goal of the technologies used is to reduce the costs of infrastructure maintenance through automated inspections, to facilitate systematic documentation of structures and to conduct risk assessments. Current limitations were also discussed and recommendations for the future given. The appendix of this report gives a detailed description of the methods and technologies behind IBM Research's visual defect detection.

Data organisation

Pixmap carried out three different collection missions with varying requirements, which was particularly relevant in terms of data volume. This allowed IBM Research to optimise the efficiency and accuracy of the collected image data. Mission 1 (M1) aimed for the highest level of detail accuracy which, however, had a direct impact on the total flight time of the drones, the evaluation time and on the scope of data processing and costs. To address challenges of handling large data volumes and to reduce costs of drone inspections utilising AI models, the requirements were relaxed in M2, with a target **GSD** (ground sampling distance) established at 0.75 mm/pixel, representing a threefold relaxation compared to M1. This reduced the data volume by a factor of nine, saving time and storage space during data processing. Additionally, IBM Research observed secondary benefits such as a reduced need for image overlap and acceptance of minor increments in motion blur. As well as further reducing the capture time and data volume, these modifications provided for an efficient and streamlined process while still ensuring sufficiently high data quality for analytical purposes. Under these circumstances, the higher risk of overlooking small defects in M2 was acceptable. Mission 3 (M3) primarily served as a comparison mission with specific overlap requirements and covered less total area. IBM Research analysed the large amount of image data (> 15,000 high-resolution images) from each mission separately, to find the optimal method for organising and processing data as regards detail, accuracy and efficiency.

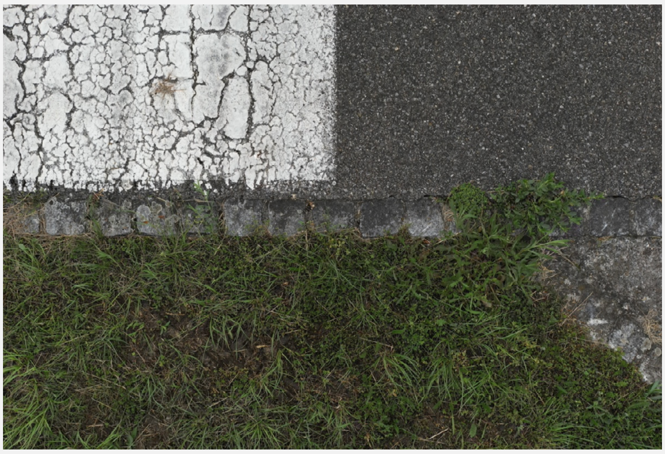


Figure 3: Original image from M1. All of the images from M1 have a resolution of 0.25 mm/pixel, which results in a very high data volume.

Image stitching algorithm

The main objective of all three missions was to merge many individual images into a large, detailed and accurate overview image showing the entire section of the runway. The challenge lay in the extreme precision required to detect cracks and defects in detail.

IBM Research developed a special algorithm that operates in multiple phases to merge the images efficiently and accurately. The method focussed on minimising positioning and alignment errors to create a coherent and precise overall image. To effectively process the immense volume of data, the dataset was divided into smaller segments that were processed independently and merged later.

Key aspects of image composition included:

Extreme accuracy: the GPS system used by pixmap gmbh had an accuracy of up to one centimetre to optimise the positioning of individual images. This corresponds to a much higher level of detail compared to conventional methods, making the image composition very demanding.

Complex algorithms: the developed algorithm worked in multiple stages to optimise, align and ultimately merge the data into one large image.

Data management: in order to manage the immense quantity of data, it was broken down into smaller, manageable parts that were later merged together again.

For the sake of processing efficiency, methods were implemented to save and reuse already processed data, thus saving valuable computing time.

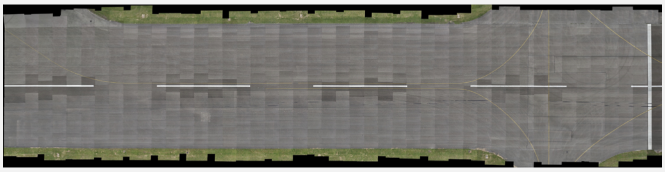


Figure 4: Overview image (M2) of the runway section based on 816 original pictures. Any spot on the overview image can be digitally inspected in detail using the zoom function.

AI foundation model technology for reliable crack detection

IBM Research's **foundation model** developed for crack detection is based on an automated visual inspection data model obtained from a broad range of civil infrastructures. The AI uses vision transformers and self-supervised learning. It goes through several stages: initially, it is trained on general images, then on already available images of civil infrastructures such as bridges, to subsequently be refined for the specific task of detecting cracks. This model was applied to the images from missions M1 and M2 and was able to deliver reliable results, despite noisy detections at the edges between the kerb and the actual runway surface. The automated detection of cracks in images is reliable with this model. The data from M2 proved sufficient for the analysis and detection of relevant cracks and offered a good balance of quality and effort, making it feasible to capture the entire runway in half a day.

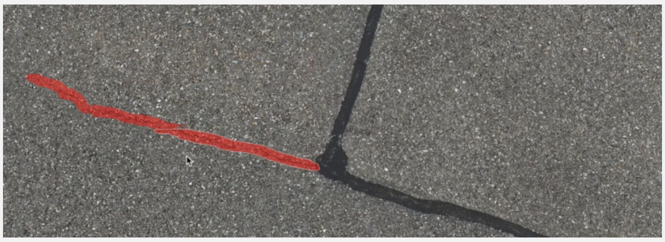


Figure 5: The automatically detected cracks are highlighted in red and supplemented with information on length and severity of the defect (see next section).

Mission	M1	M2	M2 vs. M1
Total crack instances	3,920	2,629	32.9% less
Confidence (>0.5)	691	586	15.2% less

Table 2: Overview of number of detected crack instances for missions M1 and M2. The AI model was run in a sensitive mode, thus detecting many cracks. The numbers shown in the second line of the table are lower as only crack instances with a confidence of at least 0.5 were considered. The difference between M1 and M2 is proportionally much smaller here, which means that the model delivers good results even under more challenging conditions.

IBM OCL tool for presenting results

The IBM One-Click-Learning (OCL) platform is a research instrument developed to present and display AI results in a comprehensive manner. It is used to visually represent, generate, demonstrate and iterate AI results. In this project, the image and prediction viewers were used to display cracks and defects on the runway in Dubendorf.

Features and views of the OCL tool:

Image viewer: Allows access to and navigation of all provided images (> 15,000) in structured folders, focussing on the three different missions and image qualities.

Prediction viewer: Displays the results of the AI models and automatically extracts associated attributes such as crack length to represent pixel-accurate segmentation masks of defects in the runway.

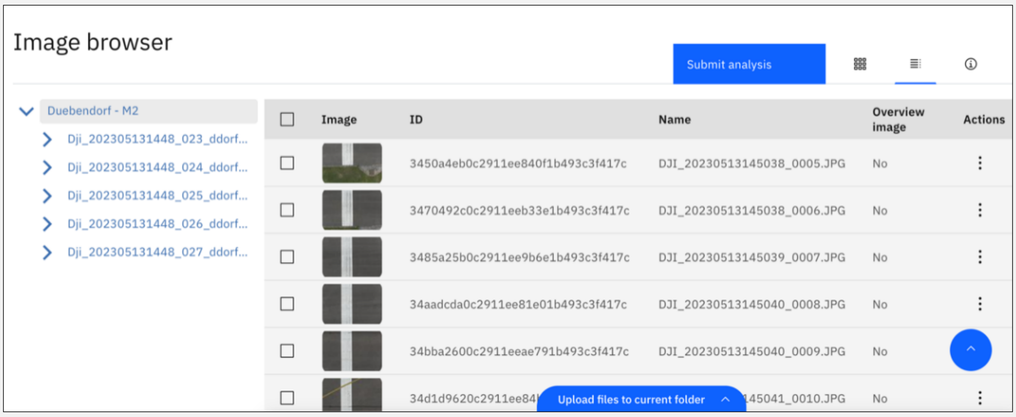
Overview viewer: Allows the user to view large sections of the runway and navigate in real time. It also enables understanding of the context in which a defect was detected.

Merged predictions and overview: Enables the user to understand the relationships and positions of all defects and consolidates multiple detections of a defect in different images into a single prediction.

Statistical viewer: Aggregated statistics are provided on all detected defects, both for individual images and for the entire merged overview.

Reporting functionality: Reports can be created for all data captured in the tool, with detailed views of each defect, important attributes and direct links to OCL for an enlarged view of the visualised defect. Georeferencing allows runway maintenance staff to manually check defects on the runway and repair damages with a sealant.

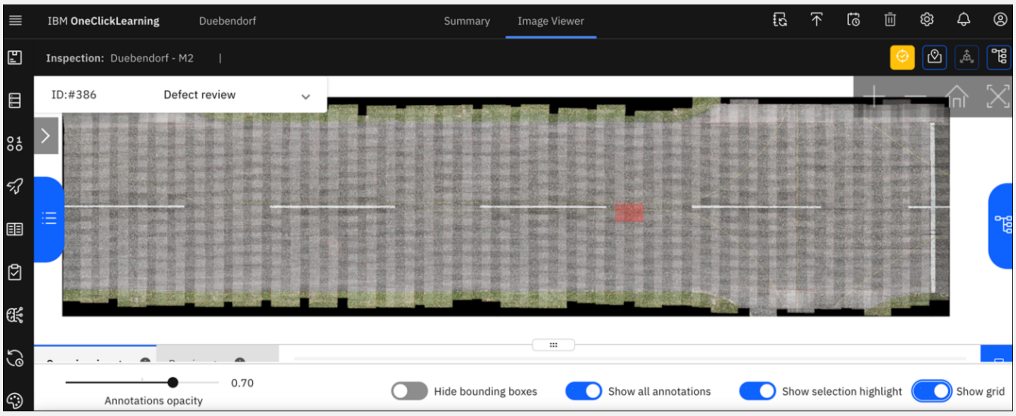
In the project, OCL was primarily used to demonstrate and verify results, particularly in the context of missions M1 and M2, to facilitate the analysis and interpretation of large volumes of data. IBM Research provided a comprehensive report and specific crack detections for the M2 data. This is especially relevant for the use of the project results by the competent infrastructure operators.



Duebendorf - M2
- Dji_202305131448_023_ddorf...
- Dji_202305131448_024_ddorf...
- Dji_202305131448_025_ddorf...
- Dji_202305131448_026_ddorf...
- Dji_202305131448_027_ddorf...

□	Image	ID	Name	Overview image	Actions
□		3450a4eb0c2911ee840f1b493c3f417c	DJI_20230513145038_0005.JPG	No	⋮
□		3470492c0c2911eeb33e1b493c3f417c	DJI_20230513145038_0006.JPG	No	⋮
□		3485a25b0c2911ee9b6e1b493c3f417c	DJI_20230513145039_0007.JPG	No	⋮
□		34aadcda0c2911ee81e01b493c3f417c	DJI_20230513145040_0008.JPG	No	⋮
□		34bba2600c2911eeae791b493c3f417c	DJI_20230513145040_0009.JPG	No	
□		34d1d9620c2911ee84	45041_0010.JPG	No	⋮

Figure 6: OCL image viewer allows access to all original project data.



Inspection: Duebendorf - M2

ID:#386 Defect review

Annotations opacity 0.70

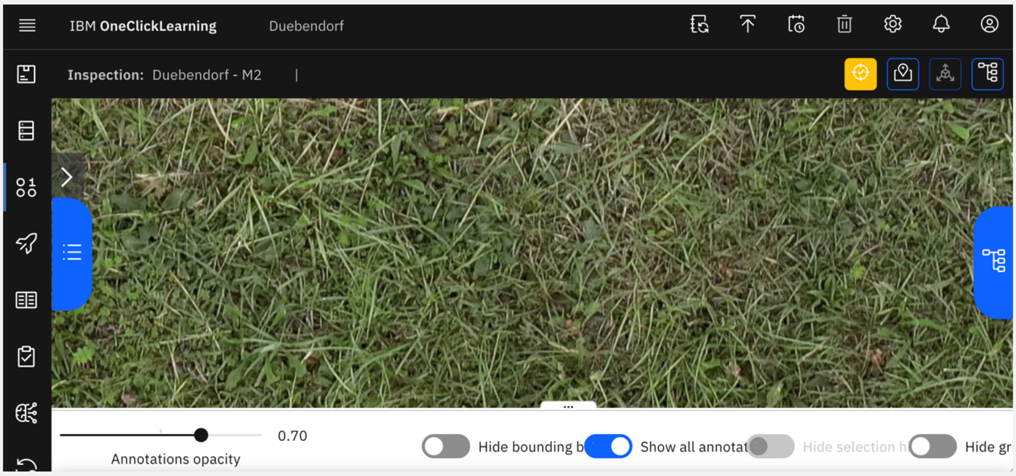
○ Hide bounding boxes

● Show all annotations

● Show selection highlight

● Show grid

Figure 7: OCL overview viewer allows users to view large sections of the runway in full resolution.



Inspection: Duebendorf - M2

Annotations opacity 0.70

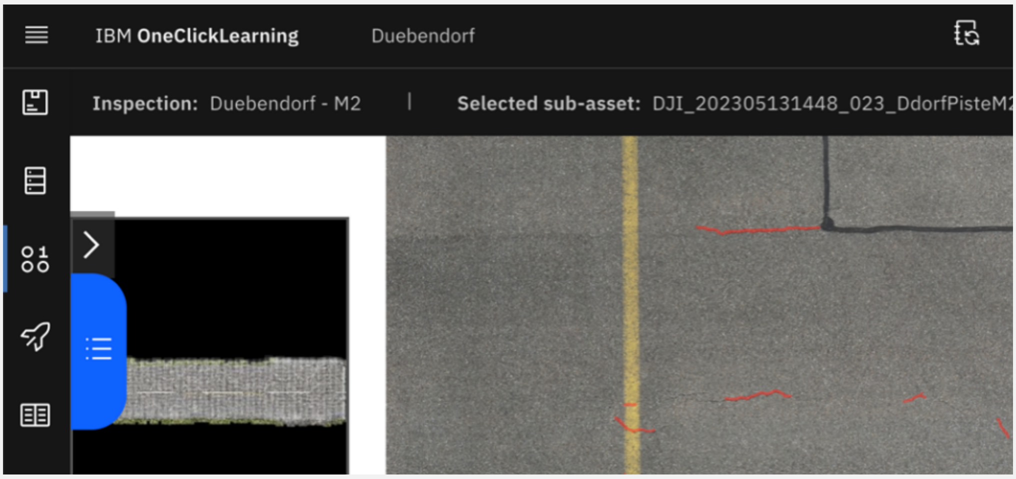
○ Hide bounding b

● Show all annotat

○ Hide selection h

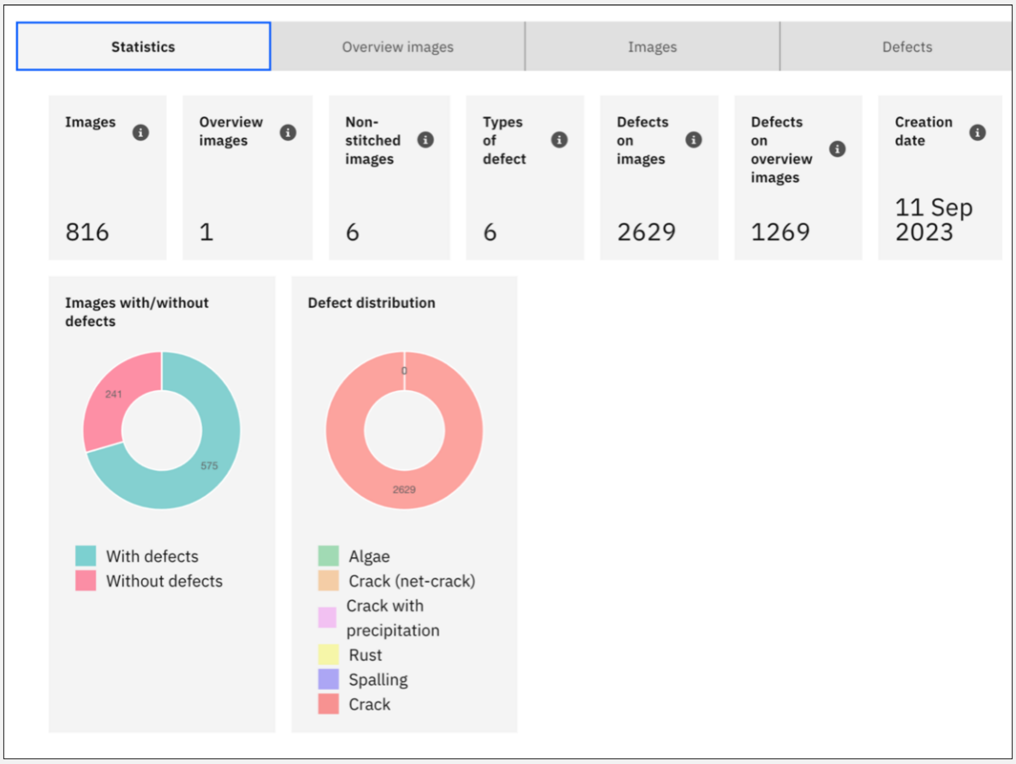
○ Hide gr

Figure 8: OCL overview viewer which fully preserves the details in full resolution.



Inspection: Duebendorf - M2 | Selected sub-asset: DJI_202305131448_023_DdorfPisteM2

Figure 9: OCL image viewer on which crack detections are marked in red on the original image.



Images	Overview images	Non-stitched images	Types of defect	Defects on images	Defects on overview images	Creation date
816	1	6	6	2629	1269	11 Sep 2023

Figure 10: Statistical viewer of M2 data on OCL

IV.

Potential of AI in infrastructure maintenance

The Innovation Sandbox project **“Automated Infrastructure Maintenance - Drone Inspections with Computer Vision”** has demonstrated successfully that large volumes of data can be collected systematically and evaluated by AI models. Cracks were identified correctly in all three missions (M1, M2 and M3). In practice, to ensure that large areas can be scanned within a reasonable time, the collection process must be carried out efficiently. IBM Research has shown that M1 and M2 are superior to M3. Furthermore, one of the key findings states that the resolution of M2 is sufficient to detect the relevant cracks, to provide full documentation on them and to make sound decisions with respect to the overall condition of the infrastructure. The entire runway can be captured by a drone in half a day which, in practice, allows for continual inspecting (e.g. semi-annually in spring and autumn, to capture seasonal differences on a continuous basis).

Access to high-resolution image data of airport runways like the one in Dubendorf is usually difficult to obtain. Being able to use real-world data within the scope of the Innovation Sandbox for AI is, therefore, of great significance for the Zurich Metropolitan Area as a location for research and innovation. Institutions like IBM Research are thus given the opportunity to evaluate and improve the latest AI algorithms and strategies in a relevant context. Therefore, this type of data will also contribute to advancing future developments of AI technology in the domain of automated image recognition in the years to come.

Furthermore, every project in which AI applications are successfully used adds to the assurance that the developed foundation models –in this case by IBM Research– operate reliably in a broad context. This means that this type of image recognition can also be used for inspecting facades of large buildings, bridges, dams, tunnels or road surfaces.

Confirmation of added value of automated runway inspections

This project has confirmed that image recognition offers great potential for the automated inspection of infrastructure elements. Even if the AI application is not yet in operational use, it can be assumed that it will be possible to address the four core challenges of manual inspection mentioned in chapter one.

Greater efficiency

The project has shown that use of drones to collect imagery saves time. Furthermore, greater efficiency can be achieved through use of image recognition, compared to traditional inspection methods performed by ground staff without AI-based reporting. Even when AI technologies are in use, an expert will, in most cases, need to conduct a final on-site validation and evaluation of any detected defects. However, the ground staff can carry out the inspection with the aid of an existing decision-making basis. Ideally, the reports with the largest defects would be forwarded directly to the maintenance company responsible for repairing the damage. This would optimise and expedite the entire process.

Better documentation

The opportunity to create digital images of infrastructure elements and to continuously check them offers considerable added value. At present, complete, objective and uniform documentation is, in practice, often lacking. Digital imagery is especially useful for detailed and systematic checking. It promotes quality assurance given that existing cracks and already repaired defects can be monitored accurately and over a longer period of time.

Fewer sources of human error

Use of automated systems helps to minimise human errors which often occur due to differing and inconsistent assessments by experts. Especially in the event of changes of staff or short-handedness, traditional inspection methods may lead to varying evaluations. Automation of inspections through AI allows for a consistent and objective analysis and evaluation of infrastructure elements, thus rendering the results more reliable and comprehensible.

Greater safety

In dangerous environments, such as dams or bridges, drones can be used to perform inspections in order to minimise the risks for humans. Although this aspect was not the focus of the current project, it is an important aspect, particularly considering that inspectors have to expose themselves to potential hazards in such environments. Through use of image recognition technology in such environments the safety aspect can be increased while at the same time providing a detailed and precise analysis of the respective structures.

Outlook

The objective of the Innovation Sandbox for AI is to strengthen the innovation ecosystem of the Zurich Metropolitan Area. The project at hand is contributing to this objective. However, widespread use of image recognition in infrastructure maintenance is still a distant prospect. Therefore, to make even better future use of the potential of Zurich as a location for innovation, the project team proposes the following action points:

1. Integration of automated inspection in existing processes

With a view to maximising the added value of automated inspections, it is important to integrate these technologies seamlessly into existing processes of infrastructure operators. This entails, inter alia, the development of interfaces in order to interlink various applications and to provide the results in a format that allows for further processing by infrastructure operators. Especially during site inspections on foot, the maintenance staff need access to a digital model of the runway with GPS function in order to find and verify the cracks identified through automated crack detection on site. Furthermore, best practices should be made available across organisational boundaries so as to share knowledge and experience effectively and to thus facilitate the introduction and use of automated inspection technologies.

2. New open data approaches within the innovation ecosystem

In order to strengthen the innovation ecosystem in the Zurich Metropolitan Area, it is essential for new open data approaches to be developed. To that end, more data based on specific use cases from industry, research and public administration should be made available. This will enable other stakeholders within the innovation ecosystem to implement similar projects. The provision and use of large volumes of data from airport runways, bridges and dams will allow for innovative solutions to be developed and the potential of AI-based inspections made better use of.

3. Transferability to other infrastructure elements

Transferability of the tested methods and technologies to other infrastructure elements such as bridges, roads and dams needs to be explored. Every use case comes with specific opportunities and challenges, which is why interdisciplinary dialogue and collaborative thinking across various infrastructure categories is important. Such dialogue will promote the development of adapted solutions for diverse infrastructure elements, thus enabling broad use of innovative image recognition technologies in infrastructure maintenance.

4. Proactive handling of regulatory questions

Overcoming regulatory obstacles is crucial for implementing AI-based inspections with drones. Important measures in this regard include setting up test environments for experimental applications. In-depth conversations with regulators, including the Federal Office of Civil Aviation (FOCA), and industry associations (e.g. Drone Industry Association Switzerland) are important too, in order to define requirements early on, adapt laws in a future-focussed manner and to speed up certification processes. These strategies can help to reduce innovation barriers and promote the introduction of new inspection technologies in the Zurich Metropolitan Area as well as Switzerland-wide.

5. Strengthening social acceptance of drones

A key factor for the successful implementation of AI-based drone technology for infrastructure maintenance is the strengthening of social acceptance of drones. It is important to have an open and inclusive dialogue with the general public, so as to communicate opportunities and risks transparently, and to clarify misunderstandings. Information events, workshops and educational initiatives could help increase knowledge about drone technology. Furthermore, meeting zones and experience parks could be set up where people are given the opportunity to come up close with drones, experience their features and applications and, by so doing, gain a better understanding of drone technology. These measures could contribute to reducing prejudices and be conducive to promoting trust in drone technology, which, in turn, will support the implementation of innovative inspection technologies.

V.

Appendix: Technical IBM Research Report

Recent research trends in deep learning are enabling more and more specific applications in computer vision tasks that would have been unimaginable a decade ago. Two main factors are fuelling this development: firstly, the broad availability of annotated data allows research to develop, test and improve algorithms, and, secondly, the availability of high-end computer infrastructure, including dedicated hardware such as graphical processing units (GPU), is accelerating experimentation times and enabling training, testing and the implementation of larger AI models.

Within this context, automated visual defect detection in civil infrastructure (e.g. bridges, buildings, or runways) becomes realistic, the goal being to use automated inspection to reduce maintenance costs, to enable a systematic way to document infrastructure at scale and to perform an adequate risk assessment. Machine learning (ML) techniques that learn from a vast amount of annotated data are trained in a supervised manner, so to speak, to gain insights that allow them to make predictions on new, previously unseen data. Deep learning (DL) techniques are a subset of machine learning algorithms that have achieved breakthroughs in many computer vision applications, including image classification, image semantic segmentation, image instance segmentation and video classification.

Typical image vision datasets available in research are extremely large - for example, ImageNet features more than 1,000,000 annotated images with class labels available for each image, which is a prerequisite for supervised deep learning algorithms. By contrast, the task at hand of civil infrastructure defect detection involved a few challenging fundamental aspects: firstly, the amount of publicly available relevant annotated data in this area is small; secondly, visible defects collected systematically in existing datasets are rare; thirdly, fine-grained instance segmentation annotations either do not exist or are challenging and time-consuming to create.

Deep learning-based computer vision research has shifted the focus to address these fundamental issues. Few-shot learning is used when only few annotations are available. Transfer learning trains AI models in one domain and applies them to a slightly shifted but similar domain to study generalisation behaviour. The latest trend of foundation models is moving towards training even larger and more generalist models from the start which are used either directly or indirectly in downstream tasks. Self-supervised learning (SSL) techniques are used as they allow for images to be recognised and reconstructed during the pre-training stage without having to rely on annotated data. That being said, for this method to be used successfully, pre-training will need to be followed by supervised fine-tuning with some annotated image data relevant to the target task.

IBM Research Zurich has extensive expertise working with client-specific data in the domain of visual civil infrastructure inspections, most notably in bridge pillar inspections, but also roadway/asphalt inspections, radio tower inspections and general defect detecting in urban areas, including walls and building facades. The project goal within the Innovation Sandbox for AI was to demonstrate the end-to-end flow and the actual value of automated visual inspections. Additionally, attention was to be drawn to existing limitations and future recommendations made as to how the technology can be further improved.

The main contributions of IBM Research to the AI Innovation Sandbox project are listed below:

- Comparison of three collection methods (missions 1 to 3)
- Image stitching technology to accurately handle low overlap setups
- Demonstration of AI models to detect cracks
- Presentation of all of the results using the OCL tool
- Short video summarising the results and on use of the OCL tool
- Reporting of all findings in a PDF document

Organising large amounts of data

Three data collection missions were carried out by pixmap gmbh (see chapter 2), to obtain information about the effects on flight time and to compare the image recognition results. IBM Research processed the image data of all three missions independently of each other in order to gain conclusive insights for the variants and to allow for a direct comparison of the data collection methods.

Mission 1 (M1) had a conservative resolution requirement of 0.25 mm/pixel which was achieved by pixmap gmbh in all captured image data of M1. These very strict requirements had direct implications on the overall flight time, the overall evaluation time and on the captured data volume. Data volume notably influences how efficiently data can be stored, processed, and analysed. To tackle these three challenging and costly factors, IBM Research relaxed the GSD requirements in M2. GSD stands for ground sampling distance and describes the distance between two consecutive pixels measured on the ground. The target GSD of 0.75 mm/pixel for M2 represented a three-fold relaxation compared to M1. The effects were multifarious: the relaxed requirements had an exponential effect on the expected data volume per unit area reducing it by a factor of nine. Secondary effects, such as less overlap and acceptance of the risk of a slightly more motion blur, reduced the data capturing time and data volume. However, any deviation from the standard recommended settings harboured the risk of a higher error rate in the process of producing the overview images, and/or of differing results obtained in the AI analysis.

In particular, the risk of missing tiny defects was higher in M2 given that the theoretically smallest detectable defect in M2 is increased three-fold compared to M1. That notwithstanding, M2 proved

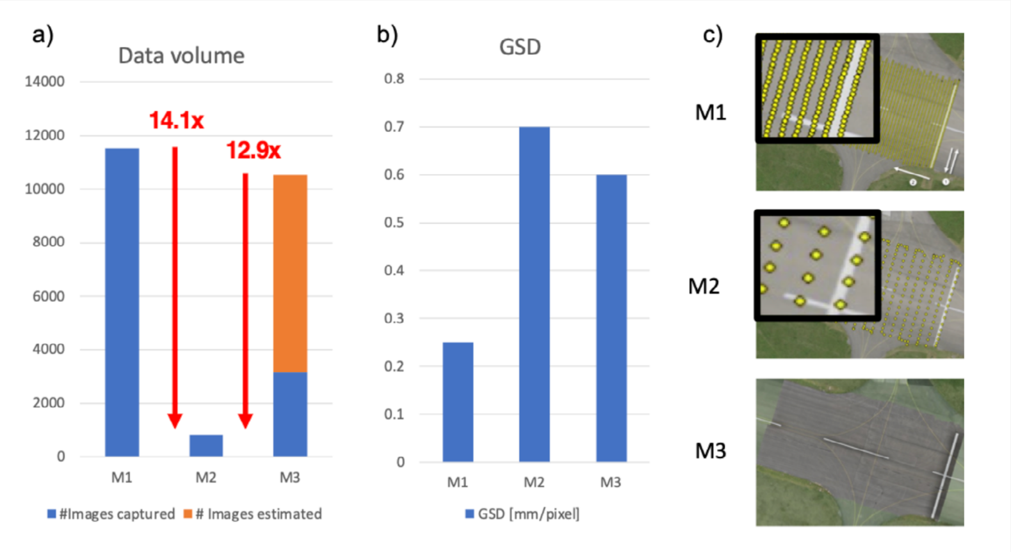


Figure 1: Overview of the captured missions M1, M2, and M3. Subfigure a) shows the number of captured images. Subfigure b) shows the ground sampling distance per mission. Subfigure c) shows a preview of how the data was captured. M3 was carried out on just 30% of the total area; therefore, the estimated number of additional images which would be required to cover the same area of 200x40m as in M1 and M2 was added.

to be a viable solution that is ten times faster and able to cover the same area as M1. M3 was carried out with large overlap requirements that corresponded to those of external tools. pixmap gmbh additionally provided the georeferenced M3 orthophoto as a comparison. M3 covered only roughly 30% of the total test area that was fully covered in M1 and in the fast M2 setup.

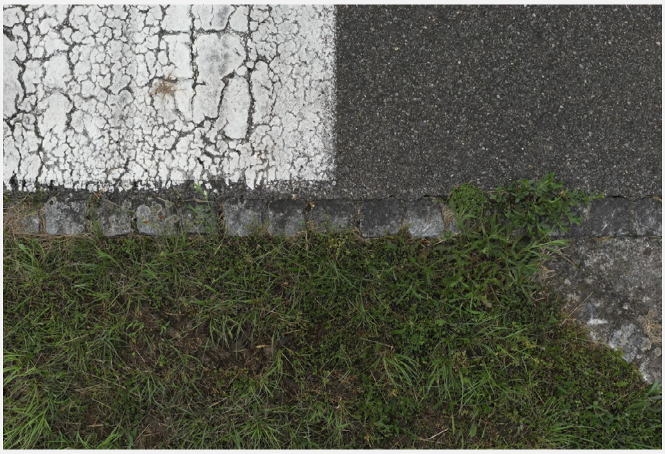


Figure 2: Representative sample of the originally captured images of M1. The image resolution is 0.25 mm/pixel.

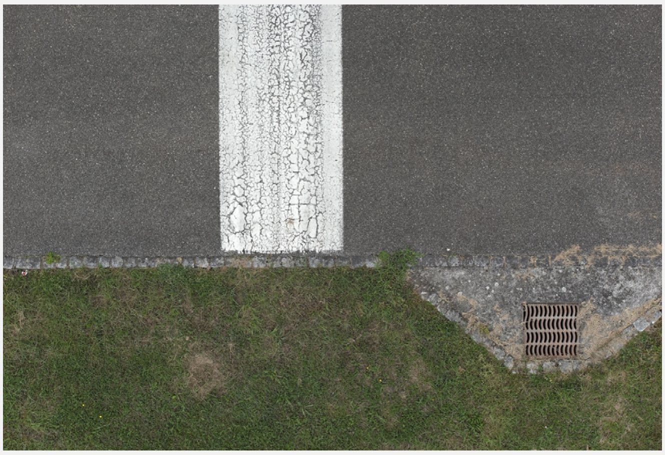


Figure 3: Representative sample of the originally captured images of M2. The image resolution is 0.75 mm/pixel. The GSD of M2 is three times larger than that of M1. This becomes apparent from the approximately nine times larger area covered by the M2 reference image compared to M1.

Image stitching algorithm

Image stitching algorithms typically require high overlaps of about 70% to work reliably. Furthermore, drone missions for mapping are usually subject to less stringent requirements, with accuracy of a few centimetres. The impact of these factors was that the image data for this project was captured at the limit of what is considered technologically feasible. In addition, the stricter the demands, the greater the effect of regular noise in physical systems, e.g. imprecisions of GPS location, deviations from the planned route due to external conditions (wind), and camera-related errors (e.g. deviations from the expected orientation). However, consistent identification of individual cracks requires a pixel- and sub-millimetre accurate location alignment. By contrast, regular GPS location receivers only deliver metre-accurate positioning under moderate assumptions. In this project, pixmap gmbh used a real-time kinematic (RTK) enabled GPS that delivers a nominal accuracy of one centimetre. Typically, alignment requirements exceed this quantity so drone location information is considered to be precise and triangulation can rely heavily on the drone location. By contrast, missions M1 and M2 had shorter GSDs and, hence, stricter alignment requirements below the noise level. This atypical assumption rendered the image reconstruction challenging for M1 and M2.

For this project, IBM Research developed and applied a customised variant of image stitching technology to solve the described challenges. IBM Research's algorithm variant operates in three main steps. In a first step, information is extracted to define a global optimisation problem. In a second step, the optimisation problem is solved and, in a third step, all data is post-processed to produce all tiles allowing to populate the data that is displayed by the viewer. The first step is further decomposed into a neighbourhood search of close images based on the GPS metadata to enhance the processing time. For each established pair, a two-step image stitching problem is solved to visually align those image instances. Corresponding point pairs from both images are extracted for the matched key points passing the filtering stage. The two clusters of key points must match each other in the final view, henceforth defining error equations related to a single pair. Pair equations are extracted for the collection of established pairs. Due to the planar nature of the data at hand, the final optimisation problem is formulated directly in the 2D image space of the target overview image that is assembled.

The loss function consists of three terms: visual alignment, size regularisation and position regularisation. The visual alignment loss is computed over the pair equations aligning key points for each two images. Size regularisation is introduced over the length of the projected image edges ensuring they match a reference length of the original image size. Without this loss, the optimisation problem tends to shrink images as, while it does not improve the alignment, it still consistently improves all error equations. Position regularisation is introduced so that all projected centre points of the optimised image positions follow the same pattern as recorded within the GPS position attached to each image. As presented above, the nominal accuracy of the GPS is limited. Therefore, a relatively weak regularisation factor was used to account for potential errors present in that information. Still, the position regularisation loss helps to place images with global accuracy on the positions they belong. Without this step, the algorithm tends to seamlessly solve the alignment (so that pairwise overlaps among images look good), but also has the tendency of errors accumulating in such a way that a drift pattern becomes visible when images are not globally placed correctly. Accounting for all three loss factors, IBM Research was able to process M1 and M2 data to reliably produce overview results for the entire scene.

The reconstruction for M2 data was done directly for the entire scene consisting of 816 images. n*(n-1)/2 pairs exist for n images yielding a total of 332,520 image pairs. Applying the clustering to the GPS data and only considering image pairs where the drone was positioned in five-metre proximity, the number of considered pairs was reduced to 19,023 pairs, thus reducing the number of equations by 94.3%.

M1 consisted of a vast amount of data that exceeded 10,000 files, making it challenging to process everything all at once. With a view to treating the data more efficiently, reducing the work per task, potentially benefitting greatly from parallel task processing, allowing for partial recovery and shortening iteration cycles during debugging, IBM Research decided to split the data into groups that were treated independently. This divide-and-conquer principle showed that it made sense to keep a limit of maximum 200 images per optimisation run. To also maintain the project structure, which provided for delivery of M1 data in 14 blocks, IBM research further decided to subdivide each block into five data chunks. Hence, seventy independent problems, each consisting of 164.7 images on average, were solved, following pre-processing, global optimisation solving and post-processing as described above.

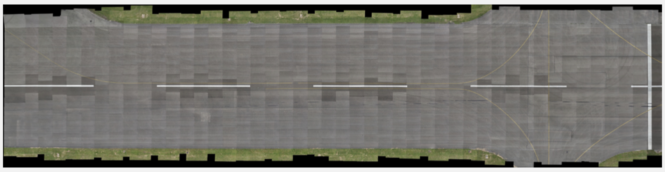


Figure 4: Preview of the overview of M2. Even in the challenging setup with little overlap, the stitching algorithm was able to fully reconstruct the full scene consisting of 816 original images. The test area encompasses 200 metres of runway and is 40 metres wide. The overview image is available with a GSD of 0.75 mm/pixel.

To further process data efficiently, IBM Research implemented merging routines that took the results of multiple solved subproblems and solved the implied larger problem. This process was implemented efficiently, so that work already performed in a subproblem could be stored and reused. For instance, extraction of all interest points and their feature vectors were not recomputed but reloaded, thus saving a substantial amount of computing time. In addition, error equations for the merging problems were only updated for image pairs that connected new image chunk pair equations. Image pairs already present in existing chunks were reloaded. The algorithm was able to process all M1 data in each of the seventy chunks without any problem. IBM Research performed the merging routine 14 times in order to merge all five sub-chunks into each block as data was provided.

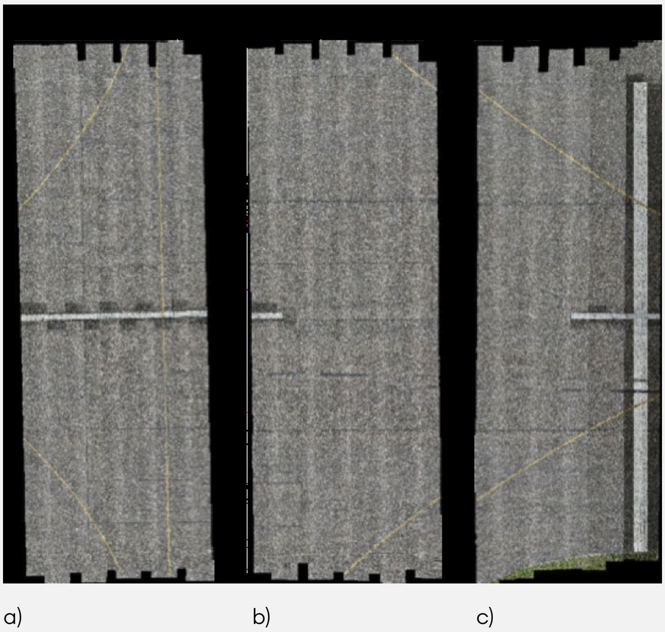


Figure 5: Preview of the M1 merged stitching results per block. Subfigures a), b, and c) each show the reconstructed overview corresponding to one block as data was provided. Result images are available in full M1 resolution of 0.25 mm/pixel. On average, each of the 14 merged blocks consists of 823.6 original M1 images.

AI technology for reliable crack detection

The AI model used for crack detection is a result of a foundation model for visual inspection which IBM designed to tackle detection tasks on general civil infrastructure.

The model uses the recent vision transformer architecture, which relies on self-attention blocks rather than traditional convolutional blocks to extract information from images. This foundation model for visual inspection is trained in a hierarchical manner using IBM's foundation model pre-training pipeline. The different stages of this pipeline iteratively tailor the model towards a particular task all while using mostly unlabelled data, i.e. data containing no annotations of defects.

In the first stage of the pipeline, the model is trained on large-scale common object datasets using self-supervised learning (SSL). This consists of solving a pretext task, such as reconstructing an image after having obfuscated parts of it. This teaches the model to process the image data and attend to specific parts of it so that it can do its task successfully. The result is a model that understands and can now extract useful information from images.

The above step is repeated once again. However, in the second stage of the pipeline, a more technical dataset is used. In this case, the model uses a civil infrastructure dataset. This is a collection of images of bridges, roads, and so forth. The result of this stage of the pipeline is a model which is now tailored towards civil infrastructure. In other words, the model now understands civil infrastructure images and can process them properly.

The final step involves fine-tuning the model using supervised learning to perform the final intended task, also known as the downstream task. The model is trained using labelled images and is instructed to locate and detect pixels belonging to certain defects, in this case detecting cracks. This training step involves using images with cracks already labelled on the images. The model's task is to locate the same cracks given only the images to start with.

The result is a foundation model for visual inspection which is particularly tuned towards detecting cracks on civil infrastructure images.

Summary of project-specific results

IBM Research ran the AI model described above on all original images from missions M1 and M2. M3 was examined by injecting an externally assembled image. Since the full scene was subdivided into tiles of 1024x1024 pixels in size, the main analysis of this data was carried out on the said tiles. IBM Research considered M3 data as additional data demonstrating an external assembly, but that only covers 30% of the total test area. Therefore, the AI team focussed on comparison of M1 and M2 data. Note needs to be taken of the fact that both M1 and M2 are superior to M3 in terms of capturing volume/quality ratio, given that M1 has a roughly three times better GSD for similar data volumes, and M2 has 12.9 times less data than M3 with a similar GSD. M1 and M2 missions achieved far better quality within roughly the same time, or delivered the same quality with less effort involved. Accordingly, the results of M2 are considered sufficiently good for a fair assessment of the section of the runway under study. Furthermore, if requirements are met, it is feasible to capture the entire runway (approx. 2.8 km) in less than a day.

Mission	**M1**	**M2**	**M2 vs. M1**
Total crack instances	3,920	2,629	32.9% less
Confidence (>0,5)	691	586	15.2% less

Table 2: Overview of number of detected crack instances for missions M1 and M2. The total consists of all detected instances, which is quite large since the AI model was run in a sensitive mode with all predictions with confidence above 0.2 reported. Note that the second line reports lower numbers where instances with a confidence of at least 0.5 were considered. Even though the GSD of M2 is three times inferior to that of M1, only around 33% fewer defects (or 15%, respectively) were reported.

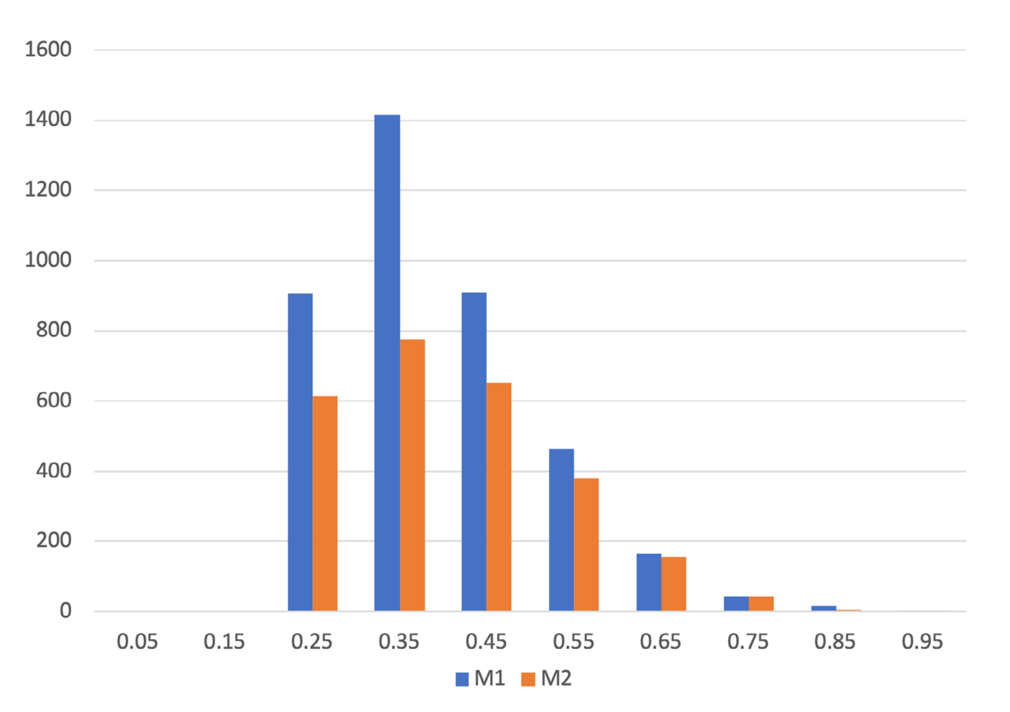


Figure 6: Frequency of AI model predicted confidence scores computed among all results of the M1 and M2 missions. Results below a score of 0.2 are clipped away. Both results follow a similar distribution; however, M1 tends to predict more instances than predicted on the M2 image material. This difference is larger towards the peak of the distributions. However, towards the confident tail of the distribution, more similar results are obtained, indicating that the M2 setup is well suited to provide similar results to M1, especially if the main results are concluded from a filtered view of confident crack instances.

Next, IBM Research performed a qualitative test where the same crack instance was selected through the OCL annotation viewer to demonstrate that all three views based on the three missions M1, M2, and M3 provided similar results. All models tended to make false crack predictions at the edge between the kerb and the actual surface of the runway. The noisy detections explain the noise components present in the AI detections. However, the real crack, which is of significance since it extends into the actual surface of the runway, is reliably detected in all three missions. In cases like the one presented, that are large enough to be visible on the image (even if captured at the lower resolution of M2), the AI models of IBM Research are able to detect them. In this sense, the M2 setup is sufficient to detect cracks similar to the one shown and to fulfil the needs of the overall analysis.

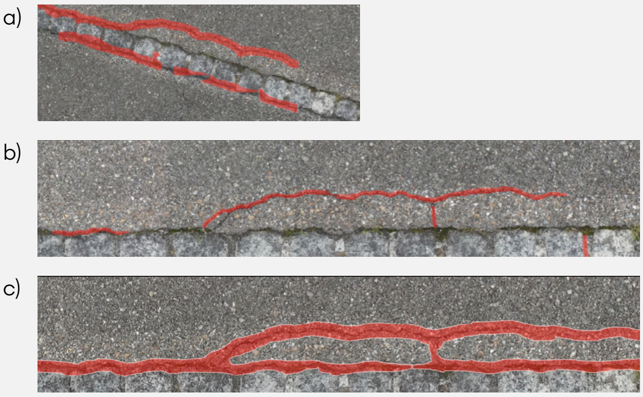


Figure 7: Predicted defect viewed through OCL. Comparing zoomed views focussing on the same location of the runway. Subfigure a) shows M3 results which are aligned so that north is at the top; hence, the image is rotated. M1 and M2 are aligned so that the main axis of the runway corresponds to the image orientation. Subfigure b) shows results for M1 (fine setup) and subfigure c) for M2 (fast setup). The real crack, which extends into the surface of the runway, is reliably detected in all three setups.

OCL tool for presenting results

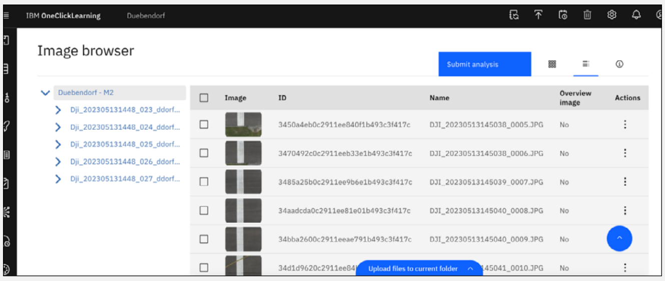
The one-click-learning (OCL) tool is an IBM research asset continuously developed to demonstrate, populate, produce, iterate and view AI results. For the scope of this project, only share and showcase functionality of the tool was used.

Main OCL features:

- **Image viewer** to access images
- **Prediction viewer** to access all AI model results for each image
- **Overview image viewer** to access the overview of the runway
- **Overview** containing merged view of merged predictions
- **Statistical view** summarising predicted defects
- **Reporting functionality** exporting documentation of defects

Image viewer

The image viewer allows access to all images provided. The tool can keep any nested folder structure to navigate the vast volume of provided data. The M1 mission was structured into 14 blocks covering the runway from east to west, the M2 mission was structured in five blocks covering the runway in the same direction. No raw images were populated for M3 since the tool is used to demonstrate the full flow when injecting an externally assembled orthophoto.


IBM OneClickLearning
Duebendorf
Image browser
Submit analysis
Duebendorf - M2
Dji_202305131448_023_ddorf...
Dji_202305131448_024_ddorf...
Dji_202305131448_025_ddorf...
Dji_202305131448_026_ddorf...
Dji_202305131448_027_ddorf...
Upload files to current folder

☐	Image	ID	Name	Overview image	Actions
☐		3450a4eb0c2911ee840f1b493c3f417c	DJI_20230513145038_0005.JPG	No	⋮
☐		3470492c0c2911eeb33e1b493c3f417c	DJI_20230513145038_0006.JPG	No	⋮
☐		3485a25b0c2911ee9b6e1b493c3f417c	DJI_20230513145039_0007.JPG	No	⋮
☐		34aadcda0c2911ee81e01b493c3f417c	DJI_20230513145040_0008.JPG	No	⋮
☐		34bba2600c2911eeae791b493c3f417c	DJI_20230513145040_0009.JPG	No	
☐		34d1d9620c2911ee84	45041_0010.JPG	No	⋮

Figure 8: Screenshot of the OCL image viewer which provides access to all original project data.

Overview image viewer

With this feature, scenes of any size can be viewed with full precision and almost in real time. To provide such a service on custom data, IBM Research has developed and runs front and backend services to maintain a tile server that can host and dynamically load on demand the request part that is currently displayed. For each view, data must be injected following a specific format of tiling the full view so that it can be efficiently served. For the scope of this project, IBM Research made sure to implement functionality to retrofit and inject large TIFF files into the image viewer tool, to directly support the orthophoto data provided in M3. All of the results obtained with IBM Research's variants of the image stitching algorithm, were dumped, stored and populated in order for users to view them efficiently.

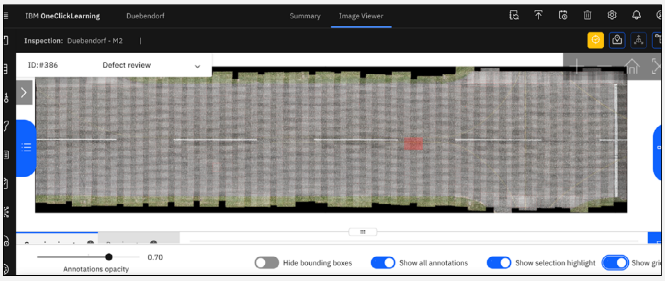


Figure 9: Screenshot of the OCL overview image viewer which allows access to the entire scene in full resolution. In this view, the M2 project shows the full test area, with the original images highlighted and the size of a selected original image displayed. This view can be used to navigate directly to the original image.

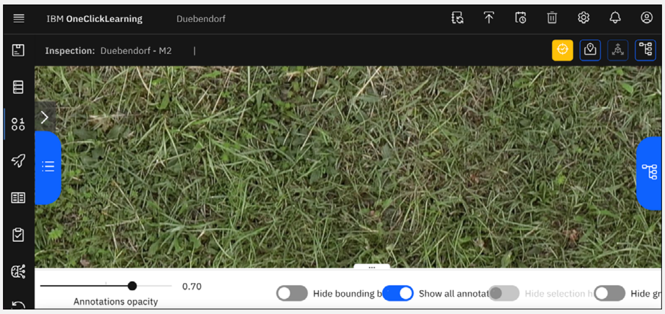


Figure 10: Screenshot of the OCL overview image viewer, with details fully preserved in full resolution. For a better understanding of the achieved resolution, grass details captured in the background of M2 are displayed.

Prediction viewer

Unless stated otherwise, AI models are always deployed elementwise on the original images, providing elementwise results. In this Innovation Sandbox project, IBM Research deployed an AI instance segmentation model to obtain pixel-accurate segmentation masks of cracks. Related attributes such as the length of the crack are automatically extracted from the predicted geometry. Additional attributes such as the predicted score are populated as well; they measure how confident the AI model is as to a certain prediction being an actual defect.

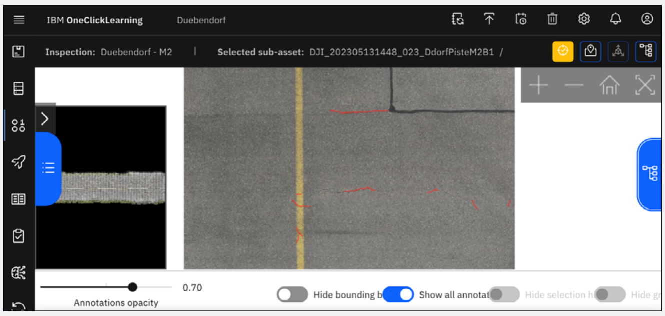


Figure 11: Screenshot of the OCL image viewer. Crack detections are marked in red on the original image. This view shows captured images and related AI model results. On the left, the mini map is available to highlight the part where the image is located.

Predictions on overview images

The main analysis was executed on the original images, processing over 11,500 images for M1. Inspecting, interpreting and drawings conclusions on such views by elements tends to be challenging as it becomes tricky for the user to capture the full context all at once. The stitched overview allows the user to view the direct context of each entry and to understand where the defect is located and how it might be related to other defects in the vicinity. Additionally, all entries were aligned and merged into logical entities. For example, a crack in the overlapping area of multiple images was captured multiple times and, therefore, typically detected multiple times, in each image in which it appeared. In the overview, those detected incidents were mapped into a single prediction. Additionally, the same process improved detection accuracy as defects had multiple chances of being detected and, if missed in some images, remained visible in the overview if detected correctly in at least one original image. For tracking purposes, the overview image can display the location of each original image as well as reference back to each original detection for every defect in

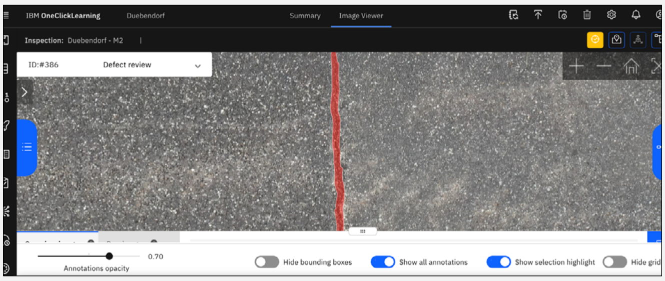


Figure 12: Screenshot of the OCL overview image. A crack detected within the scope of M2 is highlighted and displayed in the overview image viewer.

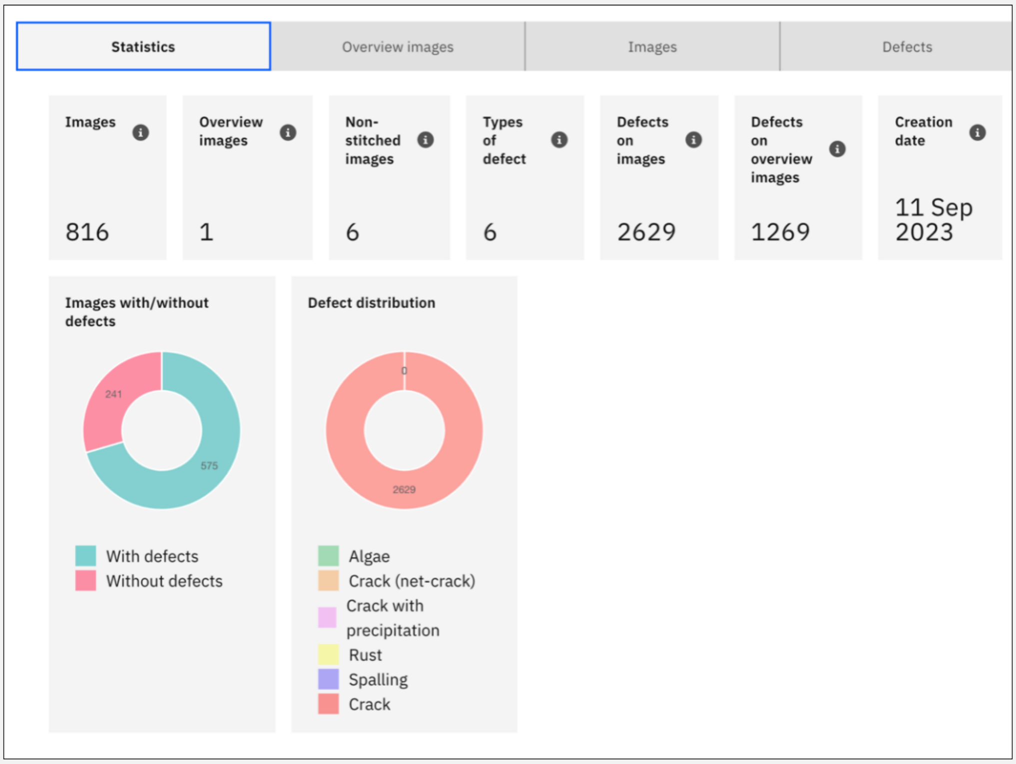
the source image material. Therefore, the overview could be used efficiently to navigate directly across all of the available data.

Statistical summary

IBM Research provided aggregated summaries containing statistics on the occurrence of all defects. The said statistics can be aggregated at each hierarchical node obtaining granular results covering only partial sections of data. The basic statistic is aggregated over the original images reflecting an elementwise analysis. In addition, this statistic was created in aggregated form over the stitched overview image, containing fewer annotations due to multiple detections being merged resulting in a single unified view of existing defects.

Reports

For all data populated in the OCL tool, reports can be extracted and produced that document all findings. The reports include an overview table section as well as a detail view section for each defect. If the tool is used to review the predicted annotations, additional information such as condition rating, defect variant classification and comments can be added by the competent expert to complete the



Images	Overview images	Non-stitched images	Types of defect	Defects on images	Defects on overview images	Creation date
816	1	6	6	2629	1269	11 Sep 2023

Figure 13: Screenshot of the OCL statistical summary page of M2 data. M2 data consists of 816 images that have been merged into one single overview image. The total number of detected defects is 2,629 in the original data and amounts to only 1,269 instances (48.2% of the original number of detections) when predictions are merged on the overview image. The majority of 70.5% (i.e. 575 images) is free of predicted defects.

report. The detail section of the report contains a detailed view of the detected defect as well as a preview of the overview depicting the actual physical location of the defect so that it can be easily found on site. Important attributes, such as length measurements, prediction scores and GPS locations, are all reported. All entries are directly linked to the OCL tool so that the hyperlink leads straight to the zoomed view of the defect visualised by the IBM Research tool.

In this project, IBM Research provided the full report of M2 data, as well as a view of selected crack detections.

Conclusion and outlook

The project successfully demonstrated how large amounts of data can be systematically captured and evaluated by AI models. Cracks were correctly identified in all three missions (M1, M2, and M3). For practical reasons, to be able to scan large areas within a reasonable time, it is imperative for capturing to be performed efficiently. IBM has demonstrated that the M1 and M2 setups are superior to the M3 setup. Moreover, the conclusion drawn is that the resolution of M2 is sufficient to capture the important cracks necessary to fully document and make solid decisions on the condition of the runway.

The data captured for the Dubendorf runway is the first dataset of its kind that provides three different mission types for the same test area. Furthermore, the inspected area was consistently captured with a resolution of 0.25 mm/pixel. Access to real-world data of this kind is extremely important for research as it allows for the latest AI algorithms and strategies to be evaluated and improved in a real-life context. As such, the data will be relevant and add value over the next decade, helping to benchmark the latest developments of AI technology in the domain of automated image detection.

Furthermore, each project within which AI is used successfully confirms that the developed foundation models work reliably in a broader context. This also increases the chances of successfully applying the same AI models to other areas, such as inspecting the facades of large buildings, bridges, dams, tunnels or road surfaces.

Glossary

Annotated data data labelled with additional information, known as annotations. For images, labels can be assigned, for instance, to objects or areas of interest. Annotated data is particularly relevant in the context of machine learning and AI.

Deep learning a subfield of machine learning based on artificial neural networks and particularly effective for recognising patterns and processing unstructured data.

Digital twin a digital representation of a physical object, process or a system that can be used for analyses, simulations and steering.

Few-shot learning a method of machine learning in which models are developed in such a way that they can get by with a very small number of examples.

Foundation model a large pre-trained model in the domain of AI that serves as a basis for building specific models. Foundation models are trained on a vast quantity of data, so as to develop a broad range of skills and can subsequently be adapted for specific tasks.

GPU (graphics processing unit) a hardware component specially designed for processing graphics and images. In AI and machine learning, GPU is used for fast processing of calculations.

GPS (Global Positioning System) a satellite-based system that provides geopositioning of objects anywhere on the Earth.

GSD (ground sampling distance) distance between two pixels measured on the ground in metres. GSD is a measurement for the resolution of an image.

Machine learning a branch of AI that allows computers to learn from data and to make decisions without being explicitly programmed.

Orthophoto aerial photograph that has been orthorectified so that it has the geometric properties of the terrain. Distortions due to perspective and height differences are corrected.

Quadcopter a type of drone that has four rotors on one level. Quadcopters are popular because of their stability and manoeuvrability and are often used for photography, videography and inspections, including for projects that involve creating high-quality imagery.

RTK (real-time kinematic) a technology to improve the accuracy of GPS systems that works in real time and is used especially in surveying.

Self-supervised learning unsupervised learning process where models learn by making predictions in a specific context. The predictions made are, however, already known.

Transfer learning a technique in machine learning that transfers knowledge developed for one task to a new, but related task.

Individuals and organisations involved in this report

Authors

Raphael von Thiessen,
Head of Innovation Sandbox for AI, Canton of Zurich

Florian Scheidegger,
Researcher, IBM Research Zurich

Reto Weiss,
CEO, pixmap gmbh

Case study provided by the Innovation Sandbox for AI

A project of IBM Research Zurich served as a case study within the Innovation Sandbox for AI. The said organisation submitted a project proposal in spring 2022. IBM Research Zurich based in Rüschlikon is IBM's European research centre and leading institute in various fields including Information Technology, Cloud and AI. The content of this report was created between June and October 2023 based on the use case **"Automated Infrastructure Maintenance - Drone Inspections with Computer Vision"**.

Imprint

Publisher
Division of Business and Economic Development,
Canton of Zurich
Metropolitan Area Zurich Association
Innovation Zurich

Project conception and coordination
Raphael von Thiessen
Location Promotion Canton of Zurich
8090 Zurich
raphael.vonthiessen@vd.zh.ch

Concept in collaboration with:
Stephanie Volz
Isabell Metzler
Patrick Arnecke

Authors
Raphael von Thiessen
Florian Scheidegger
Reto Weiss

Design
Sibylle Brodbeck, sibyllebrodbeck.ch

Publication
This report is published exclusively in digital format and in the languages German and English

Translation
Mila Myrsep, Word for Word Ltd. Liab. Co.

Project Steering
Division of Business and Economic Development,
Canton of Zurich
Statistical Office, Canton of Zurich
Division of Digital Government, Chancellery,
Canton of Zurich
Office for Economy, Canton of Schwyz
Metropolitan Area Zurich Association
ETH AI Center
Center for Information Technology, Society, and Law
(ITSL) University of Zurich