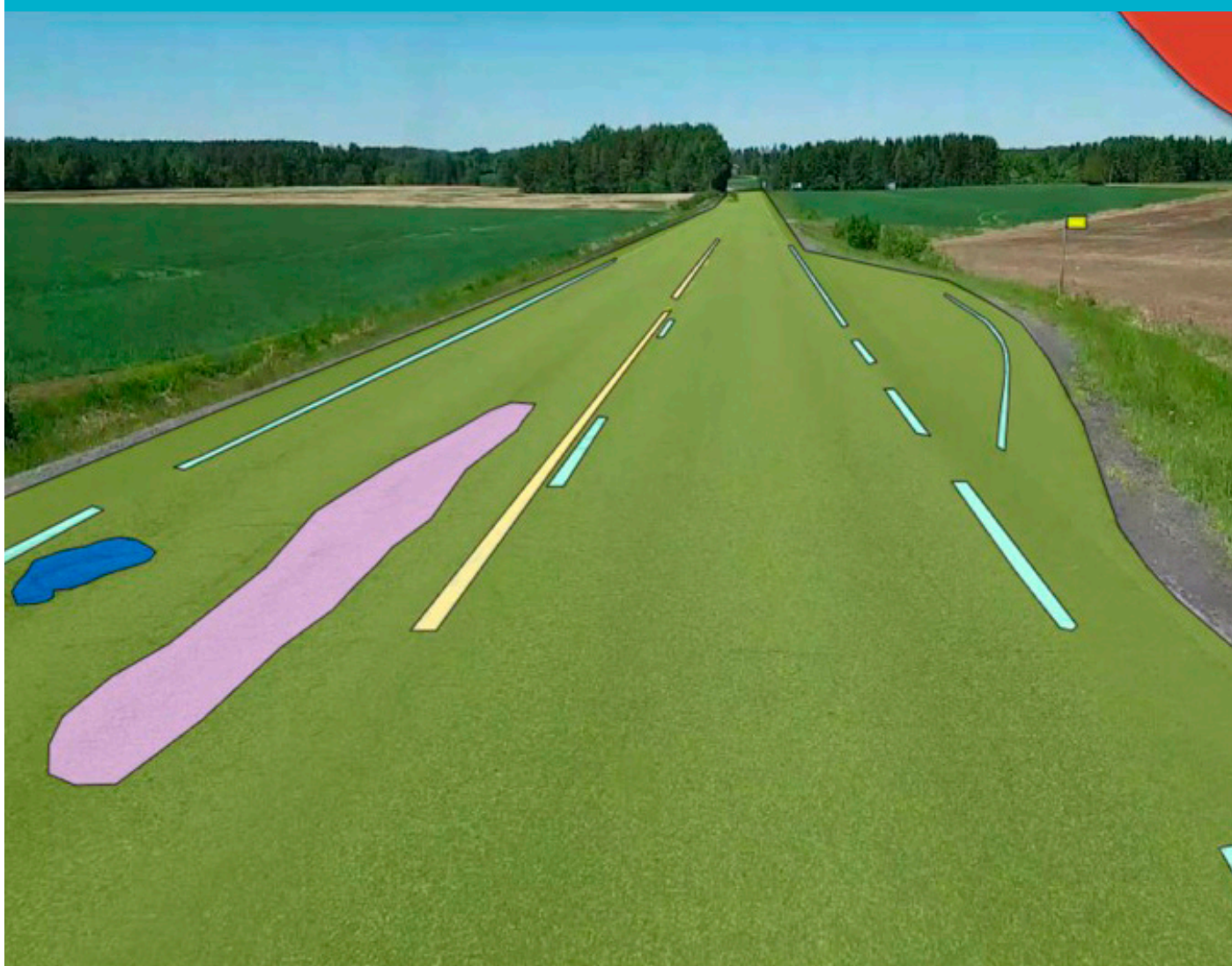# VÄYLÄ

Finnish Transport
Infrastructure Agency

DEVELOPING THE PAVEMENT REPAIR PROJECT USING COMPUTER VISION

Repair contract computer vision pilot of Pirkanmaa Centre for Economic Development, Transport and the Environment (ELY Centre)

Mikko Haavisto, Anne Valkonen, Ville Jurkkola, Antti Hannuksela

Developing the pavement repair project using computer vision

Repair contract computer vision pilot of
Pirkanmaa Centre for Economic Development,
Transport and the Environment (ELY Centre)

Publications of the FTIA 9eng/2020

Cover picture: Vaisala Oyj (picture made by RoadAI algoritm)

Abstract

Road condition management is a topical issue as road maintenance appropriations have decreased and global warming increases the weather strain, imposing even more stress on pavements when temperatures vary more and more round zero. In road maintenance, it would thus be sensible to seek to develop proactive methods that increase effectiveness and cost-efficiency.

This study is a combination of modern computer vision technology utilising machine learning and an experiment examining the development of operative contract models. The aim is to find a technical method for objectively controlling the status of holes in roads and to utilise this technology to develop a project procurement model. The study and the work carried out for it have provided good material for both the goals.

As for computer vision, it can be stated on the basis of the study that there is close correlation between the results of the human-made road damage inventory and that made using computer vision. However, a corresponding comparison could not be conducted in the interpretation of the severity of the damage. This is above all due to the fact that not even different people have a unanimous opinion about the severity of damage. In the case of computer vision, methodological development in the future should be carried as part of operational work, so that human feedback about damage severity categories can be combined into the development of automatics.

With regard to further development, research is no longer a viable point of departure for technological development. It is time to take AI applications into use alongside operational work and to develop their usability in production in the desired direction.

From the point of view of developing a repair project procurement model, the experiment suggests that computer vision can be used to determine a road condition index in the project area. However, it may be necessary at times to validate individual detections, though as a whole computer vision can produce a continuously updated hole development index, which indicates the overall condition of the road and provides status data about the success of the project.

Foreword

This study was carried out as part of the Finnish Transport Infrastructure Agency's Digitalisation Project, which seeks to enhance infrastructure maintenance by developing digital systems and operating models.

It was steered by a project team consisting of Ossi Saarinen from the Finnish Transport Infrastructure Agency; Janne Lintilä, Timo Seppänen and Tomi Keisala from Pirkanmaa Centre for Economic Development, Transport and the Environment; and Anne Valkonen from Via Blanca Oy.

Mikko Haavisto from Vaisala Oyj was in charge of writing the report, with the assistance of Ville Jurkkola from Posti Oyj and Antti Hannuksela and Markus Melander from Vaisala Oyj.

Helsinki, February 2020

Finnish Transport Infrastructure Agency
Maintenance

Contents

1 Background

The annual cost of maintaining the Finnish road network is about MEUR 500 when you take into account repairs (MEUR 212), improvements (MEUR 53) and road maintenance (MEUR 229). This sum is used to maintain Finland's state road network, whose total length is 77,995 km. Paved roads account for 50,750 km of this network (source: Finnish Transport Infrastructure Agency).

Maintaining one kilometre of the road network costs about EUR 6,200 per year. The Finnish Transport Infrastructure Agency and the Centres for Economic Development, Transport and the Environment (ELY Centres) seek to allocate this sum to the three key functions (maintenance, improvement and repair) in the most sensible and cost-effective manner possible.
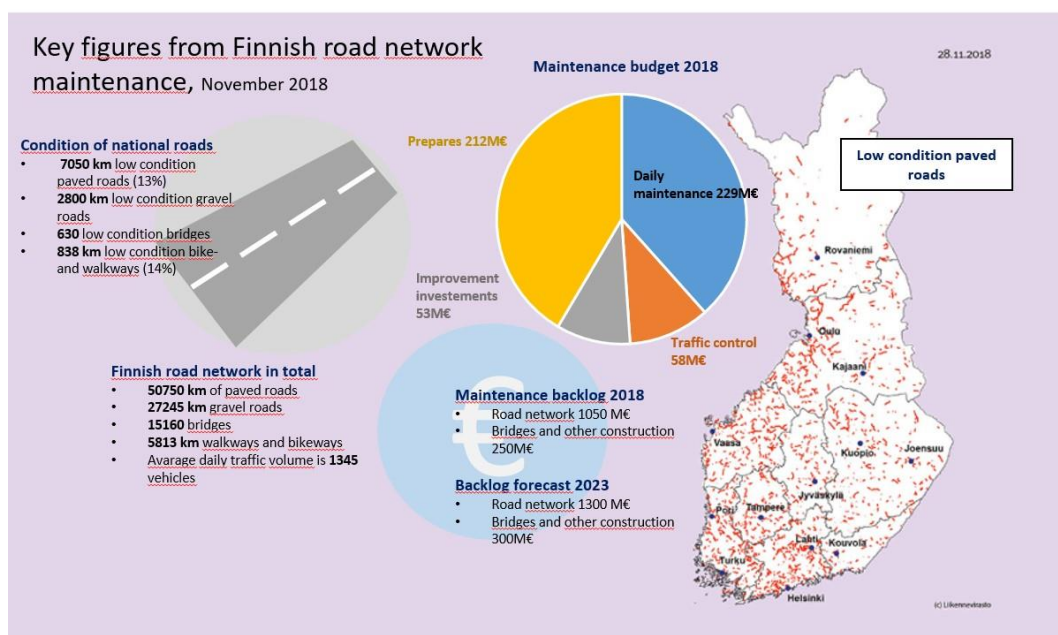
Figure 1.   An overview of road network maintenance (source: FTIA).

Pavement repair contracting plays a key role in optimising a road's service level, the surface's lifecycle, and traffic safety. Therefore, one of the most pressing challenges at the moment is to develop wearing surface contracting and find the most efficient and highest quality operating models possible.

This study was commissioned by the FTIA and Pirkanmaa Centre for Economic Development, Transport and the Environment (Pirkanmaa ELY Centre). It investigates the use of computer vision in wearing surface repair contracting, with the goal of determining how well computer vision technology is suited to inventorying and maintaining an up-to-date overview of the condition of the road network. The pilot also identified how technology could aid in the development of a procurement model for pavement repair contracting, and whether new technology could boost the efficiency of pavement repair operations and quality management in general.

The project focused on a computer vision-based condition analysis of the road network in the Pirkanmaa region, and on harnessing continuously generated data in pavement repair contracting. The objective is to develop a procurement model for pavement repair contracting on the basis of the experiences obtained during the pilot. The condition analysis project was limited to identifying holes, cracks and patches in the road network with the aid of computer vision. Contractors, clients and consultants will benefit from the data generated by computer vision and the user interface that was tested.

Two types of repairs are carried out as part of pavement repair contracting in the Pirkanmaa region: scheduled repairs and acute repairs. Scheduled repairs are based on inventory data about the condition of the road network and the need for repairs. Currently, the majority of acute repairs are commissioned via the traffic management centre.

From a technical perspective, the project focused on identifying holes that would endanger road safety, and also on detecting incipient damage that could later develop into a hole that would endanger road safety. Naturally, the ultimate goal is to arrive at a situation in which action can taken against any emerging holes at the earliest possible stage and as quickly and cost-effectively as possible.

1.1   General objectives

At workflow level, the project's objective was to pilot an operating model in which data obtained using computer vision is made available to the contractor for scheduling purposes, and in which the same data can also be used by the client for monitoring and finally as part of the contract's quality assurance process.

Piloting this operating model naturally requires the technology to function at the desired level. The first step was to ensure the desired level of repeatability in the results obtained using computer vision technology. After this, the next step will be to consider and further develop user interfaces and work processes to harness the available data in the comprehensive management of pavement repair contracting.

The following general objectives are also part of the project:

1. Encouraging innovation
   - Enabling contractors to engage in service-level thinking

2. Enabling measures to be prioritised
   - Improved forecasting and more efficient scheduling

3. Improving the quality of repairs
   - More durable repairs

4. Improving road safety
   - Reducing the amount of damage that endangers traffic

5. Using computer vision technology to obtain an improved picture of the current situation
   - Monitoring quality and measures
   - An up-to-date overview of the status of the network
   - Monitoring development

1.2 Project objectives and limitations

The project had the following primary objectives and limitations:

1. To develop a procurement and operating model for pavement repair contracts

2. To develop computer vision so that it can generate data for scheduling and monitoring pavement repair contracts.
   - The work will primarily focus on the detection of holes that endanger road safety and the identification of any incipient holes and cracks.

3. To define an imaging process that will produce a continual stream of data to support pavement repair contracting.
   - Changes to cost structures when imaging are integrated into existing processes, and are no longer carried out solely with special runs.

How or with which process should data about the current condition of the road network be collected?

How well does computer vision technology meet the data requirements of pavement repair contracting?

How could the procurement and contracting models be developed if there is more up-to-date data available?

Figure 2.   The pilot's objectives presented as research questions.

2 Developing a procurement and operating model for pavement repair contracts

In order to move from a unit price-based turn-key contract under a partnership model (KIMPPA) to a service contract model, a highly reliable parameter for measuring the condition (integrity) of road surfaces must be found. This parameter must be able to precisely describe the number of holes and cracks (that endanger road safety) and their location in the road network in a way that will enable repairs to be launched and scheduled. In the region administered by the Pirkanmaa ELY Centre, the goal is to determine whether computer vision will make it possible to switch to a model in which pavement repairs are procured on the basis of, for example, overall road surface integrity rather than individual repairs. Before transitioning to a service contract or a contractual model that contains elements of the target state operating model, it must be ascertained that the measuring parameter describes the actual condition of the road with sufficient precision.

Under a service contract model, the risk of forecasting damage in the road network would be transferred to the contractor. The service contract model is considered advantageous, as the contractor has the power to influence the durability and correct timing of repairs as well as the choice of repair method. This model would also encourage contractors to develop more efficient procedures.

Damage that endangers road safety and reduces driving comfort has been defined in the same way in the quality requirements for both regional road maintenance contracts and separate repair contracts. The specification of damage should be consistent nationwide, regardless of how the repair is to be procured.

In Finland in December 2019, there were several ongoing or soon-to-be-launched pavement repair contracts to procure repair work that had already been performed as part of regional road maintenance contracts. Some of these tasks may even be overlapping, that is, the same work is being commissioned both as part of a regional road maintenance contract and as a separate repair contract.

The client's primary target state is to ensure that damaged surfaces do not endanger the safety of road users, including pedestrians and cyclists. This is why the condition of road surfaces must be known to a sufficient degree of precision, and there must also be a way to indicate this condition with a sufficient degree of accuracy. If surface damage endangers road safety, this damage must be repaired immediately when reported, or else signage must be used to warn road users of the damage. If necessary, speed limits may be temporarily lowered. In principle, a good operating model would be one in which contractors could – and should – repair any damage that affects driving comfort as scheduled work before it becomes acute damage which endangers road safety.

A parameter to describe the condition of the road surface will therefore be required in order to develop a new, service contract-based procurement model for pavement repair contracts. It will then be possible to, for example, engage in market dialogue to determine whether contractors are ready to switch to service contracts. There has been competition for separate repair contracts in recent years. There are a number of operators and, although some contractors have specialised in one or two types of repair, sufficient tenders have been received.

Repair work can be divided into manual repairs and machine repairs. Manual repairs are typically acute repairs that are classified as endangering road safety, and are usually either service requests or reports of damage received though other channels. Machine repairs are usually scheduled repairs that fix all the damage to the entire road surface, or a section of it, at the same time.

Under a service contract model, the client would not specify the repair method, the materials, or the frequency of repairs. However, the following requirements could be set:

- repair quality, such as ridges (the difference in height between a patch and the undamaged surface)
- occupational and road safety
- the recyclability and environmental suitability of repair materials
- the amount and type of damage.

The client could also set requirements or limitations on the disturbance, dangers or damage that repair work may cause to third parties.

In a service contract, demonstrating quality is typically the responsibility of the contractor, and quality should be indicated with the aid of up-to-date status information (condition data). The contractor must also demonstrate how they have reacted to the status information. The client's quality control is based on monitoring both status information and the response thereto.

Some road surfaces are nearing the end of their useful lives, and a pavement that is full of holes is not worth repairing. Such roads should not be included in a total-price service contract, as the client should decide on repairs or other maintenance on a case-by-case basis. These roads can, of course, be integrated into a service contract via a unit-price section. Naturally, it is also possible for the roads to be resurfaced or even in some cases converted to gravel roads.

The service contract should contain change mechanisms that can be employed if the operating environment alters. The number of new paved roads varies year by year, and the number of pavings will naturally affect the extent of the network that requires repairing. There must, therefore, be a way to connect changes in the number of pavings to the contract's price formation mechanisms. During the contractual period, there may be changes in the weather or traffic volumes that affect the speed at which road surfaces become damaged irrespective of how the contractor implements the pavement repair contact. Before introducing a new procurement model, a risk division model simulation must be carried out using existing data from actual total-price and long-term service contracts.

3 Computer vision method

3.1 Description of the method

The computer vision system architecture (Figure 3) consists of three sub-areas: The collection of video and location data using a smartphone; computer vision analysis on servers; and a user interface that can generate reports, play videos and display analysis results.
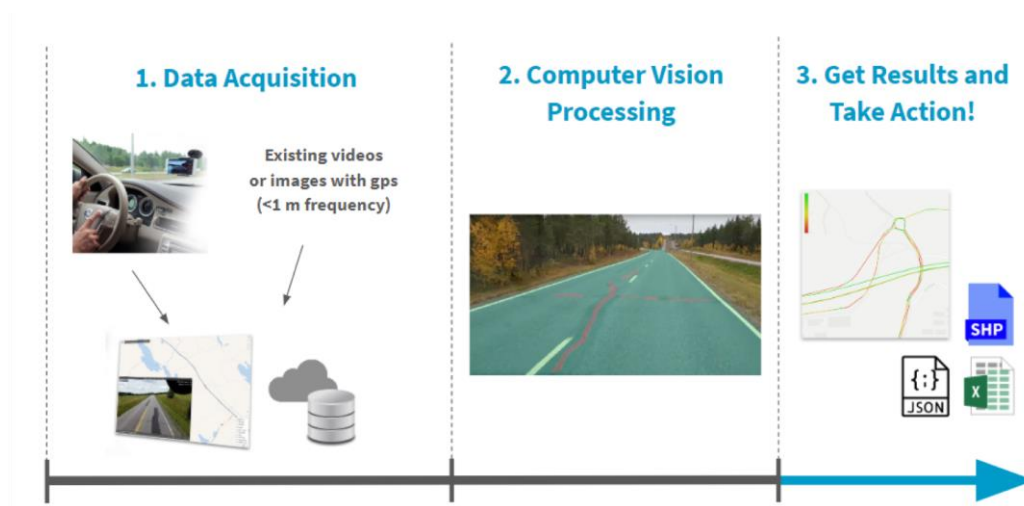

Figure 3.   Computer vision system architecture.

Computer vision analysis is based on a statistical method trained with the aid of example images. The model is then able to identify learnt categories in images that it has never seen before. In other words, the computer vision model's performance is completely dependent on the material used to train it. This training material consists of image and annotation pairs. Here, annotation refers to drawing a bounding box that contains a particular 'meaning' at image level. Examples of such meanings in a road network image could be road, lane marking, road sign, crack, hole, patch, vegetation, and so on. Figure 4 is an example of an annotated image.
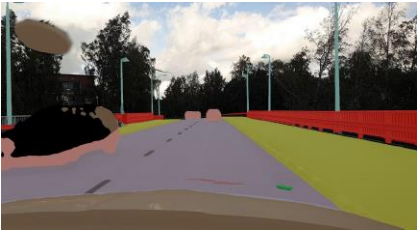
Figure 4.   *An annotated image of a road in which the colours illustrate instances of different meanings.*

3.2 Damage categories

Table 1 uses illustrations to show what kind of damage identification this project has focused on. When it comes to performance, the focus has been on identifying the kind of damage that has a critical impact on road safety. During the course of the project, the computer vision model was expanded to distinguish between more specific categories of holes and cracks in accordance with their severity (severe and moderate). Two challenges were the relatively low incidence of safety-critical damage, particularly in the severe hole category, and achieving a sufficient degree of repeatability to enable the further development of services that use computer vision to compare previous results to newer results. This could, for example, create a way to automatically report on changes in the amount of road damage.

Table 1. *A list of original and annotated images by damage category.*

Damage	Original image	Annotated image
Severe hole (green)		
Moderate hole (green)		
Severe crack (brown)		
Moderate crack (red)		
Geometric patch (violet)		
Damage-shaped patch (violet)		

The 'severe damage' categories denote damage that may endanger road safety or potentially damage vehicles. The 'moderate damage' categories denote damage that, if not repaired, will probably develop into severe damage. **In other words: severe damage should by default be repaired urgently, while moderate damage can be handled during scheduled preemptive repairs.**

An examination of the computer vision model used in the pilot indicates that, instead of using absolute limit values, the method has learnt to distinguish between the various categories of severity with the aid of the example images annotated by humans. Humans use the guiding limit values for category distinction shown in the table below.

Table 2.   Current guiding limit values for pavement repair contracts.

Type of damage	Minor	Moderate	Severe
Hole	diameter less than 5 cm, only the upper layer of the surface has broken loose	diameter 5–20 cm	diameter more than 20 cm, deep
Crack	width less than 1 cm	width 1–5 cm	width more than 5 cm

3.3 Automatic filtering of routes driven

The system used in the pilot automatically took into account variations in location data and imaging conditions when analysing road surface damage. This is an important preliminary step, so that the information entering the system can be assumed to be reliable. When preparing to integrate automated inventorying using computer vision into operative models, the following two perspectives should be taken into consideration.

It should be possible to reliably link location data to the correct road geometry and street address. The system used in the pilot produces video footage and GPS data in five-minute bursts. This information is then sent to the servers in five-minute 'chunks', to ensure integrity and make the data delivery process easier to manage. The system is programmed with an automated operating model that, if more than 75 per cent of the GPS route is on roads that fall outside the state paved road network, will filter out the raw material in question from the map user interface.

One weakness of the aforementioned automation is that if recording is not paused when transferring to the road network or to a private road, some usable inventory data may not be utilised when the results are visualised. To solve this problem, it was decided to develop a separate data model for recording information for the system used in the pilot. In the new data model, data has been saved for short sections of the road network, that is, road segments. Transferring to a segment-specific data model enables inventory imaging to be accurately paused on reaching the end of the state paved road network without the danger of losing significant data.

The damage identification system used in the pilot was developed for paved roads and functions best in good inventorying conditions, that is, on dry roads and in good light. In the target operating model, it is understood that road network imaging must be automated to the greatest possible extent and, that in a crowdsourced operating model, vehicle users will not want to pay special attention to the usability of their footage in damage identification.

Computer vision offers excellent potential to filter this kind of raw data. A computer vision-based system for identifying weather conditions has been developed on the data collection platform used in the pilot. In addition to generating data for winter maintenance, it can also be adapted to automatically filter suitable material for road surface damage identification. On a practical level, this means that any images that are unsuitable for road surface damage identification will be automatically filtered out of the technical analysis, so that no attempt is made to identify damage in the dark, for example. Figure 5 shows the categories of conditions for filtering material and recommendations for their suitability.
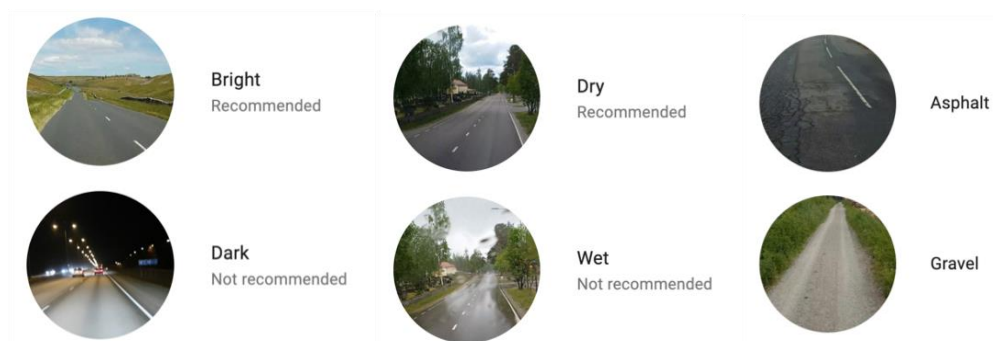
Figure 5.   Computer vision filters used in the pre-processing of raw video footage

In clear weather conditions, reflections on the windshield also pose a challenge to acquiring footage (Figure 6). To solve this, a physical polarising filter is placed over the phone's camera. The accompanying pair of images illustrate the errors in damage identification caused by reflections. Although windshield reflections can be eliminated with a simple polarisation filter, it is also important not to leave papers or other objects on the dashboard, as they may also cause reflections.

Figure 6.   An example of how objects on the dashboard can affect computer vision when no polarisation filter is placed in front of the camera lens.

3.4 Measuring damage

This section will describe how the system used in the pilot identifies damage to the road surface.

By default, computer vision identification is run for every frame of video footage. At a speed of 80 km/h and 15fps (frames per second), a series of images will be collected at 1.5-metre intervals. As identification becomes less accurate when the subject is further away from the camera, the system used for the pilot only utilised computer vision to search images for damage in a zone 2–5 metres from the front of the vehicle. Any damage detected in this zone is classified and defined by damage category, in accordance with the area/extent of the damage in relation to the width of the road. The final results of the damage analysis are stored in the database in ten-minute chunks. Each chunk is labelled with a start and end location linked to a street address.

Here, the relative width of the damage refers to the width of the damage in relation to the width of the road as measured at image level (Figure 5). This data further helps to differentiate between the different categories of damage. Damage severity is also used to filter out errors in computer vision identification: for example, less than 1 per cent of the holes detected are typically just artefacts in the image.
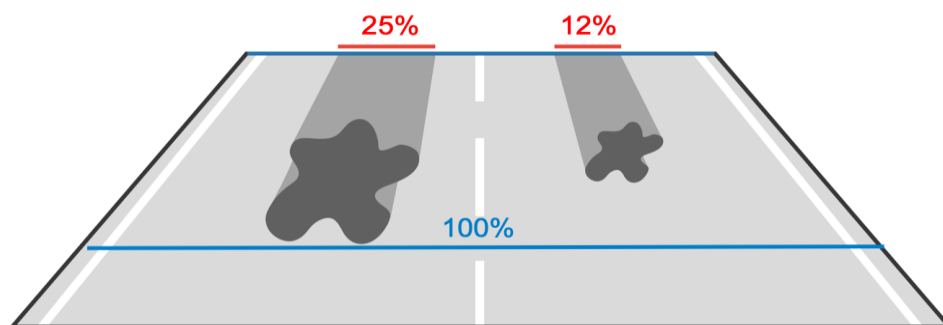
Figure 5.   Width of the damage in relation to the width of the road.

3.5 Identification zone

Damage is currently identified across the entire width of the carriageway. Pedestrian and cycle paths are not included in the identification zone and are automatically filtered out of the zone. In the future, it will be possible to limit identification to a specific lane, which will make it easier to compare width values between different roads. Repeatability can also be expected to improve if identification is limited to the lane in which the data collection vehicle is driving. As this lane is closer to the camera, any identifications made will be larger and more reliable at image level. The disadvantage of lane-specific identification is that video footage must be produced for every lane, which increases the need for data collection runs. Figure 6 shows how the map user interface marks cycle paths in yellow, indicating that the path has been correctly identified and the results from this zone will not be used in road surface damage identification.
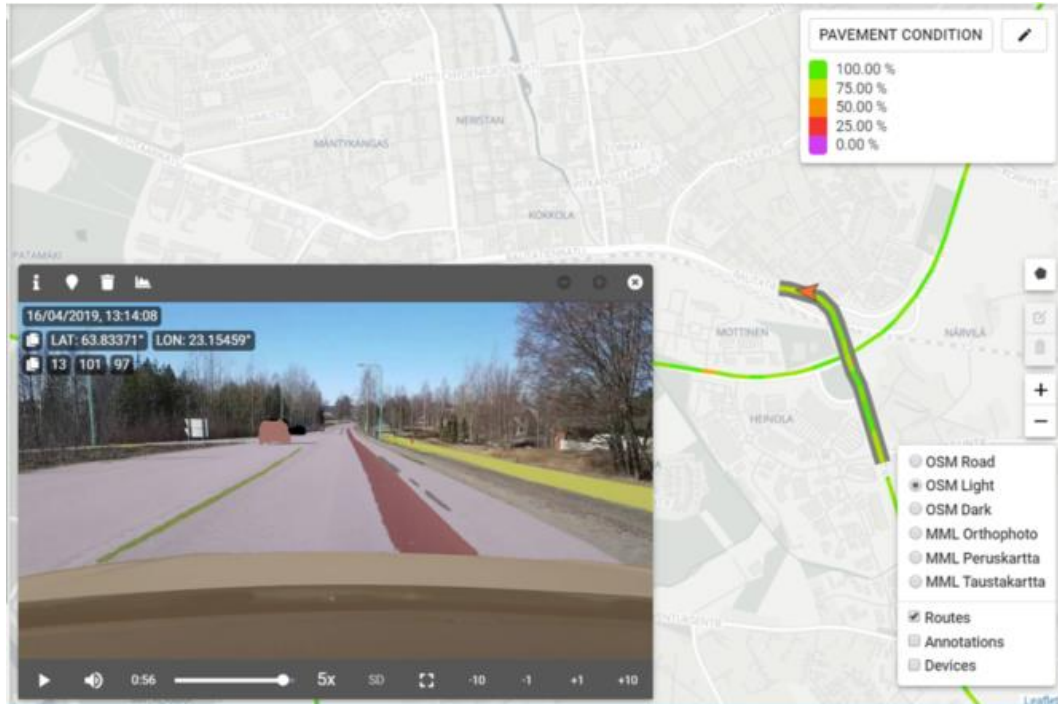
Figure 6. *The identification zone is limited in several ways. In the picture, the zone marked in yellow has been classified as a cycle path.*

4 Other uses for the video service and computer vision platform

4.1 The video service and scheduling repairs

The basic premise is that all parties involved in a pavement repair contract will have access to the latest version of the video service and the map app that visualises serious damage and patches. The app will provide both a map overview of the total amount of damage and, using the videos, the possibility to verify the need for repairs and indicate the required measures without visiting the site itself.

The impact of the repairs will be reported back to the system by filming the section of road in question after the repairs have been carried out. This means that, for example, patches will now be identified instead of holes.

To provide additional guidance in making repairs, the contractor can create a legend to label damage identified in the videos, for example, 'immediate repair' or 'to be monitored'. The system will then be able to produce a PDF report complete with map and terrain images. This PDF report will be easily distributable, either digitally or in paper format as work guidance for employees in the field.

4.2 Computer vision as a scheduling tool for paving

The material produced by the computer vision system can also be used for other purposes outside pavement repair contracts. The production cycle for footage used in a repair contract or maintenance processes will be noticeably quicker than, for example, the material gathered for scheduling paving. The same video footage could therefore be used more extensively to identify damage categories that would support the scheduling of annual paving. Video footage combined with more extensive identification of damage categories would enable measures to be planned and selected in many places without the need for separate visits to the site. This process synergy would lead to increased efficiency for road network owners compared to the current situation, in which the information for scheduling paving is produced by a number of completely separate processes.

Other areas that could potentially benefit are condition assessments of lane markings, puddle formation on roads, the identification of non-indigenous species and the need for coppicing, and the detection of deficiencies in ditch drying.

5 Data collection process for pavement repair contracts

During the project, the plan was to mainly film the road network using a smartphone installed in post office (Posti) vehicles. This was a natural choice, as Posti's routes cover about 90 per cent of the paved road network in the region administered by the Pirkanmaa ELY Centre. According to Posti, full coverage of the ELY Centre's road network would have required more than 200 separate routes. For this project, it was decided to record footage from the routes driven by eight Posti vehicles, and all material that fell outside the scope of the road network was excluded from data processing.

The basic dataset produced by the project was supplemented with material from Posti's Katudata service and one additional camera that was placed in a contractor's vehicle. One of the basic principles of the service is that the use of many data sources will enable cost-effective coverage of the network in accordance with customers' needs. This model can be used to prioritise imaging, so that roads in higher maintenance classes are covered more frequently, and the maintenance contractor's own runs can be allocated to separately selected areas that require separate data collection. From the client's perspective, one good data collection process would be for regional maintenance contractors to implement clearly designated control runs.

Footage filmed during the project:
- A total of 72,000 kilometres
- Duration about 1,410 hours

Phase 1 of the project Raw data collection and an analysis of the starting point – spring 2019

The first phase of the project in spring 2019 sought to complete a one-off analysis of the road network. Data was collected using Posti vehicles on standard delivery routes within the region administered by the Pirkanmaa ELY Centre. On the basis of the footage recorded, the average length of a run in the Pirkanmaa region was 118 kilometres. Data was collected by eight vehicles over a period of six weeks, starting on 22 April 2019. A total of 14,500 km of video footage was filmed during the spring, and it took 380 hours to collect the data. This means an average of 2,416 km and 63 hours per week. Figure 7 shows the camera transfer schedule that was proposed and implemented by Posti in order to ensure good coverage of the region.
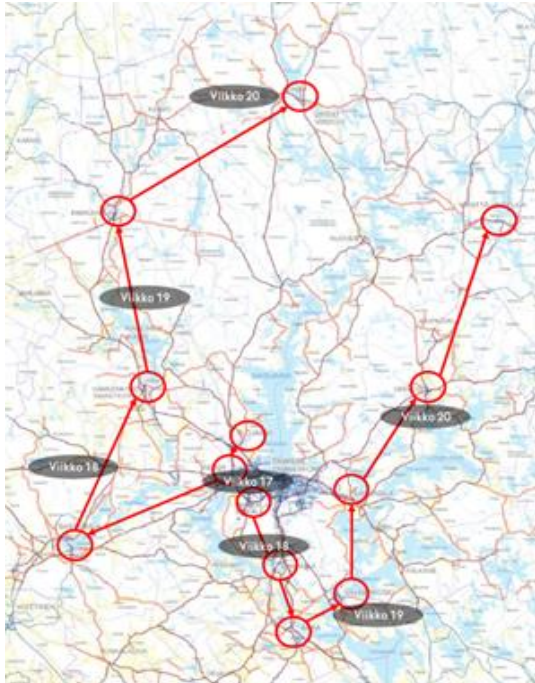
Figure 7.   Posti's plan for transferring cameras in order to cover the region.

Figure 8 illustrates the coverage of all standard deliveries made by Posti vehicles in the Pirkanmaa region. The marked routes were selected to provide the greatest possible coverage within the agreed filming period. The map shows the Pirkanmaa road network by maintenance category. The areas superimposed in black were not covered by Posti's standard deliveries.
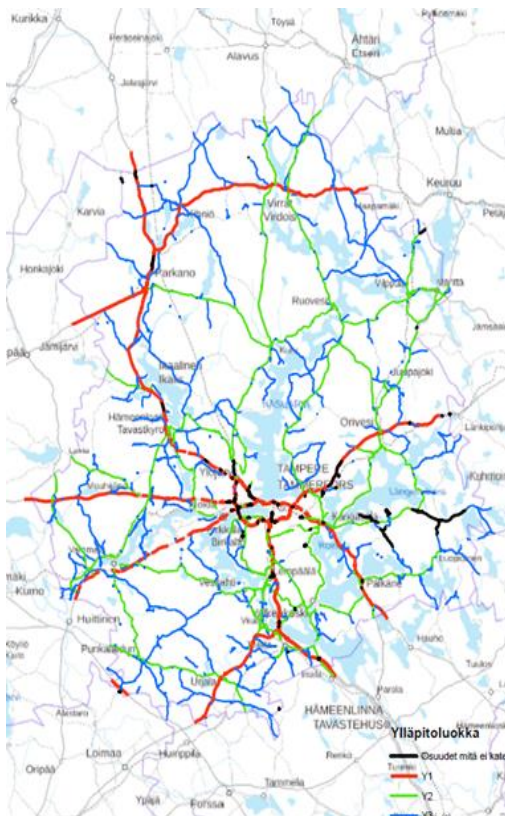
Figure 8.   The coverage of Posti's standard delivery routes in Pirkanmaa.

Figure 9 shows the imaging of Pirkanmaa's road network that was implemented during the pilot. Posti's standard delivery routes were able to achieve extremely good coverage of the paved road network in all maintenance categories. This was achieved by frequently transferring eight cameras between post office locations, and was further facilitated by a dry summer that enabled sufficient filming days. Rainy days would have caused delays in filming, which would have meant less coverage within the agreed time period. As the routes provided good coverage of all maintenance categories, we were able to film roads that were in both better and worse condition, with the assumption that the maintenance categories would generally correlate with the condition of the road.
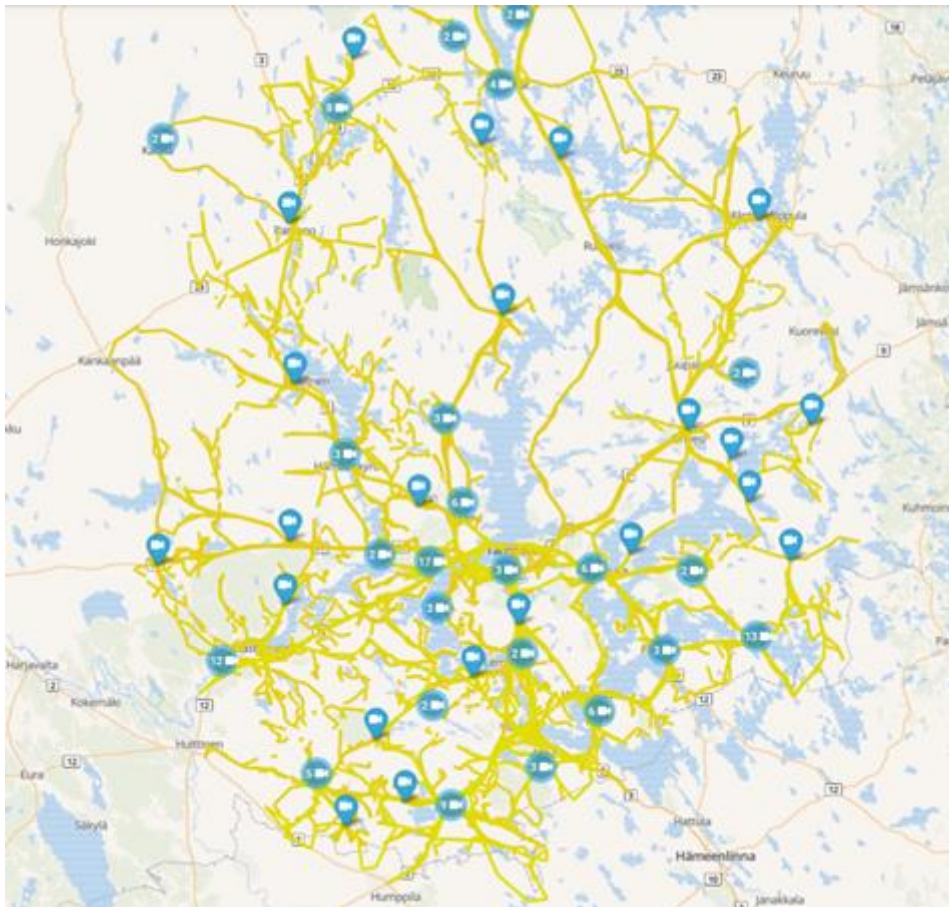
Figure 9.   Coverage of the video footage collected by Posti in spring 2019.

Phase 2 of the project   Continual data collection and repeatability simulation – autumn 2019

The next stage of the project sought to film selected road sections continuously for a period of eight weeks. Eight suitable Posti vehicles were chosen for this purpose, so that the selected road sections could be included in the inventory.

A total of 57,600 km and 1,030 hours of video footage was filmed during the second phase of the project. This equated to an average of 7,200 km and 130 hours per week. Posti divided its vehicles internally between two organisations: logistics for deliveries and daily standard postal deliveries. Data was collected from the selected routes for a period of eight weeks according to Posti's normal driving schedule, using only the selected vehicles. In addition to the Posti vehicles, one Asfalttikallio vehicle was also equipped with a camera. When

deciding on the routes and duration of Phase 2, attention was paid to the method's repeatability in order to obtain good data and experience. It was also important to obtain experience in performing condition analyses with computer vision in a variety of conditions, such as in the rain, in the dark and in bright sunshine. Figure 10 shows the coverage of phase 2.
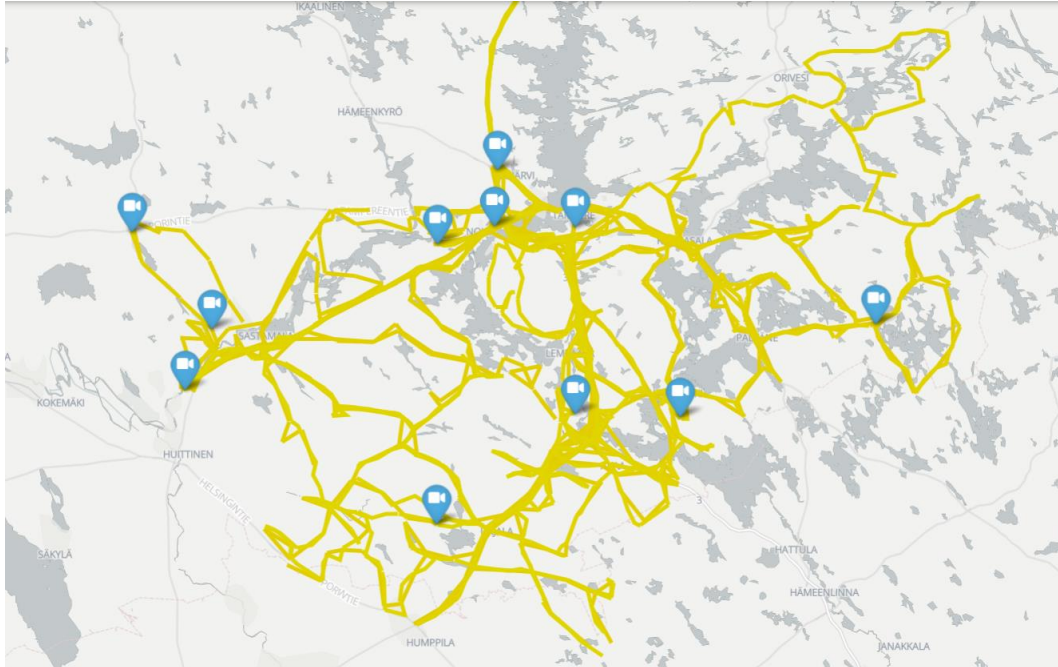
Figure 10.   Data collection coverage in phase 2.

6 Results

6.1 General technical limitations

The system used in the pilot currently records computer vision identifications with a resolution of ten metres. The relative width of damage compared to the width of the road surface is recorded for every ten metres of road. The section titled 'Further development' describes a method that could enable the number of holes to be quantified.

The closer the camera is to the damage, the more accurately the damage is identified. For example, this means that damage in the same lane as the data collection vehicle is identified more accurately than damage in an adjoining lane. This limitation also applies to the human eye, and particularly when damage is being identified from a moving vehicle.

6.2 The computer vision's performance

The computer vision's performance was assessed by comparing its results to observations made by three different people. The observations made by these people were also compared to ascertain how much human observations varied from each other. The damage noted by humans was always based on an individual frame of video, whilst the results generated using computer vision indicated damage over a ten-metre section of road. The same damage is usually visible in several consecutive images. This means that a human may mark the damage as being located at 169 m, whilst computer vision may place the same damage in the following section, at 170–180 m. For this reason, identifications are accepted as being equivalent if they are located within the same ten-metre section or in one of the adjoining sections on either side. The same deviations in location were also applied when observations made by two humans were compared.

The types of damage compared were: moderate hole, severe hole and severe crack. The results of the comparison are shown in the following table. In the comparison, the first-mentioned identification has been compared with the second-mentioned identification. The various columns show the percentages for each identification. In the table, AI denotes computer vision and 1, 2 and 3 denote the three different people who performed the annotations. For holes, all cases also indicate the percentage of identifications that were in the wrong category of severity. In the case of cracks, there is no human data available for moderate cracks. However, when the comparison is made with computer vision, the percentage of errors in identifying the severity of damage has also been indicated separately.

Table 3.   *Results of computer vision comparison.*

Moderate holes			
1: 237 detected	**Comparable identification**	Identified as a severe hole	No hole detected
1 vs 2	**0.97**	0.03	0.00
1 vs 3	**0.97**	0.01	0.02
1 vs AI	**0.93**	0.01	0.06
2: 298 detected			
2 vs 1	**0.91**	0.02	0.07
2 vs 3	**0.96**	0.01	0.04
2 vs AI	**0.87**	0.01	0.12
3: 382 detected			
3 vs 1	**0.88**	0.04	0.08
3 vs 2	**0.94**	0.02	0.04
3 vs AI	**0.87**	0.01	0.12
AI: 188 detected			
AI vs 1	**0.88**	0.08	0.04
AI vs 2	**0.90**	0.07	0.03
AI vs 3	**0.94**	0.03	0.03
Severe holes			
1: 117 detected	**Comparable identification**	Identified as a moderate hole	No hole detected
1 vs 2	**0.88**	0.11	0.01
1 vs 3	**0.65**	0.35	0.00
1 vs AI	**0.47**	0.51	0.02
2: 130 detected			
2 vs 1	**0.83**	0.15	0.02
2 vs 3	**0.54**	0.45	0.02
2 vs AI	**0.45**	0.52	0.02
3: 65 detected			
3 vs 1	**0.94**	0.06	0.00
3 vs 2	**0.86**	0.11	0.03
3 vs AI	**0.48**	0.49	0.03
AI: 45 detected			
AI vs 1	**0.73**	0.24	0.02
AI vs 2	**0.87**	0.07	0.07
AI vs 3	**0.64**	0.33	0.02

Severe cracks			
1: 47 detected	**Comparable identification**	Identified as a moderate crack (Humans did not classify any instances of damage in this category)	AI moderate crack / Human severe
1 vs 2	**0.45**		0.55
1 vs 3	**0.38**		0.62
1 vs AI	**0.43**	0.53	0.04
2: 43 detected			
2 vs 1	**0.49**		0.51
2 vs 3	**0.28**		0.72
2 vs AI	**0.40**	0.58	0.02
3: 29 detected			
3 vs 1	**0.62**		0.38
3 vs 2	**0.45**		0.55
3 vs AI	**0.38**	0.55	0.07
AI: 40 detected			
AI vs 1	**0.58**		0.43
AI vs 2	**0.48**		0.53
AI vs 3	**0.30**		0.70

Simply by looking at the amounts of different damage detected, it is evident that there are clear differences between people's definitions as well. Person one detected 237 moderate holes, person two 298, person three 382 and computer vision 188. Person one detected 117 severe holes, person two 130, person three 65 and computer vision 45. The lower number of holes detected by computer vision is at least partly explained by the fact that only one instance per damage category could be detected for each 10 metres of road, even if there were in fact more instances within the same area. Person three clearly detected more moderate holes than the others. The number of severe cracks detected varied between 29 and 47.

In general, there was slightly more consistency in the dentification of moderate holes between the various humans than between the humans and computer vision. However, computer vision was still on average about 90 per cent consistent with the humans' observations. There was greater divergence with regard to severe holes. The greatest consistency was seen between person one and person two, who detected almost the same number of instances. Person three and computer vision were more likely to define these instances as moderate holes. However, very few severe holes went undetected.

There was quite a clear variation in the number of severe cracks detected, and there is no straightforward explanation for this. When it comes to this category, it was clear that what constitutes a severe crack is very subjective. The humans interpreted the severity of the damage in different ways. In the comparison of the humans' and computer vision's results for severe cracks, computer vision's results for moderate cracks are also given. Therefore, the final column in the table for severe cracks shows those situations in which a human identified the

damage as a severe crack, whilst computer vision identified it as a moderate crack.

The following images (Figures 11, 12, 13, 14, 15 and 16) show examples of situations in which person one and computer vision interpreted the damage in different ways.
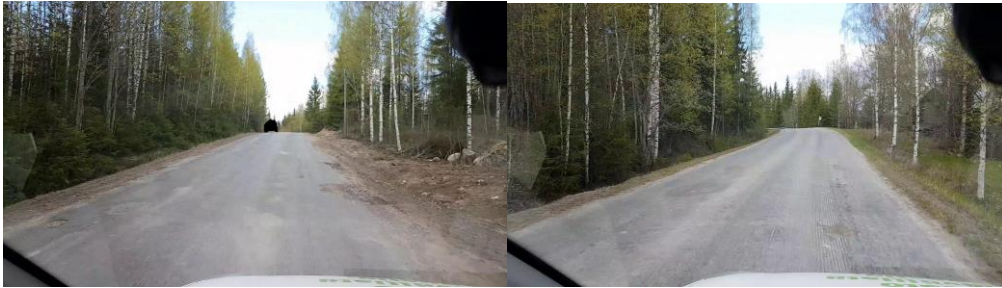
Figure 11.   Person 1: moderate hole. Computer vision: no hole identified (2 instances)

Figure 12.   Person 1: severe hole. Computer vision: no hole identified (2 instances).

Figure 13.   Person 1: severe crack. Computer vision: no crack identified (2 instances).

Figure 14.   *Computer vision: moderate hole. Person 1: no hole detected (2 instances, marked in green in the segmentation image).*
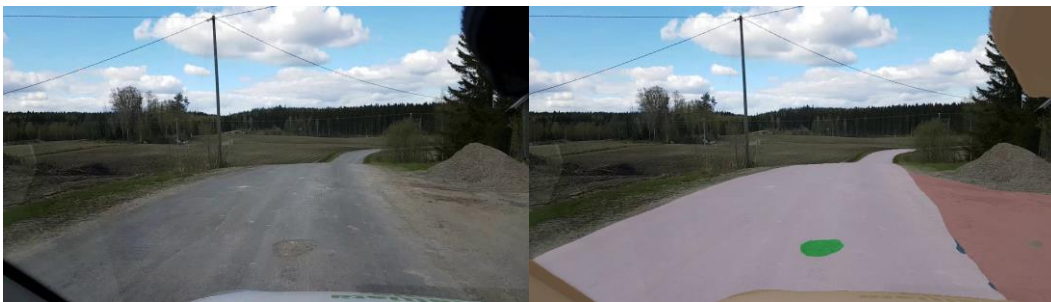
Figure 15.   *Computer vision: severe hole. Person 1: no hole detected (1 instance, marked in green in the segmentation image).*

Figure 16.   *Computer vision: severe crack. Person 1: no crack detected (2 instances, marked in brown in the segmentation image).*

On the basis of the results, computer vision identification of holes and cracks is at a good level, and the variations in accuracy are largely centred on differentiating between its severity (severe or moderate). On the basis of the comparison, the difference between the identifications made by the humans and computer vision was as follows for the various types of damage:

- Computer vision left an average of 2.3 per cent of severe holes undetected compared to the number of severe holes detected by three humans.
- Computer vision left an average of 10 per cent of moderate holes undetected compared to the three people.
- In the case of cracks, it did not seem worthwhile making a comparison or reaching any conclusions, as the humans' results differed so much from each other.

Future development of the method will focus on differentiating between the severity of damage, so that severity data can be used as an aid in prioritising pavement repair contracts. When evaluating the results, it is worth noting that the human comparison figures were observations made in an office using an observation tool, rather than whilst out driving on the road itself. If a comparison of computer vision and human inventorying in the field is desired, a short field comparison should be conducted as a separate study.

6.3 Data production requirements

The following inventory coverage requirements and summary reporting cycles were specified for roads by maintenance category.

- Maintenance class 1: 4x per month
- Maintenance class 2: 2x per month (excluding winter)
- Maintenance class 3: 1x per month (excluding winter)

The amount of footage filmed during the autumn is more equivalent to actual production. The eight vehicles filmed for an average of 130 hours on weekdays, producing 7,200 km of road videos per week. This equates to about 3 hours and 15 mins per vehicle per day. At an annual level, capacity of about 50 TB would be required to store this footage. When considering the operating model described here as a continual service, it is worth noting that, in practice, it is unlikely that all of the footage would need to be stored. On the basis of maintenance category requirements, the amount of overlapping footage could be reduced for repeated routes. This would reduce actual annual storage requirements by an estimated 60–80 per cent, which would make data storage costs more practical financially. There would be room to store reports and monitoring statistics for the entire contractual period.

Posti's data production is based on the number of regional routes and is therefore also fully equatable to the number of vehicles. Posti's coverage of the Pirkanmaa road network during the analysis performed in November 2019 was 3,300 km, with the full extent of the road network totalling 3,700 km. This equated to more than 200 distribution routes.

Posti's proposal was based on the best coverage for 50 routes. They were selected by maintenance category, with Y1 roads being the highest priority and Y3 the lowest. According to the table below, the Y3 category could be most efficiently covered using 38 Posti vehicles, giving a coverage of 491 km. These routes were driven by Posti at least once per week. The same 38 routes would cover both Y2 and Y3 roads. Table 4 below shows the most important information about the number of different road categories, Posti's overall coverage, and the coverage of the 50 routes.

Table 4.   Coverage of the distribution routes of Posti's 50 vehicles

	Pirkanmaa ELY Centre's paved roads	Posti's coverage	Coverage of 50 vehicles	Vehicle distribution per road category
Total (km)	3,659	3,305	2,203	50
Y1	872	672	491	38
Y2	1,221	1,141	699	43
Y3	1,566	1,492	1,013	50

The retention period for reports generated from the footage is contract-specific. Reports must be retained for at least three years, in case any claims for damages are made by road users. If requested, the contractor and the contract supervisor must be able to supply a statement about the condition of the road and any repairs made to it over the past three years. The video footage should be retained as per the contract, but the minimum period could be say three months, during which time the client and the contractor will have discussed the condition status parameters and any changes in the road's condition during site meetings.

6.4 Segment-specific data model

Change monitoring could be carried out at either damage or network level. The former requires unique identification of the damage being monitored, both within individual videos and between footage produced during different runs. This would pose an extremely challenging problem with regard to holes, as holes can appear as close as ten centimetres apart. If several holes appear within a ten-metre area, it is realistic to assume that there is currently no reliable method to monitor changes in these holes. In practice, the solution in this case would be narrowed to an analysis of individual observations presented at different points in time.

The problem could be significantly simplified by performing the analysis at network level. Instead of detecting individual holes, a predefined road segment would be examined. Here, 'road segment' refers to a section of road of a certain length (for example, roads could be divided into ten-metre segments). Road segments would in turn enable the creation of metrics for monitoring the condition of the network. Road geometry can be divided into road segments as per the desired resolution, for example, so that damage data is recorded at an

accuracy of 10 metres or 100 metres. The choice of resolution would be influenced by many practical factors, such as the measurement method's resolution (image density), the scalability of the storage method (database) and the difference in GPS/Glonass/Galileo accuracy between different runs. Reducing the resolution would, in practice, cause the solution to approach instance-level monitoring. Increasing the resolution would alleviate the technical challenges associated with instance-level monitoring.

Perhaps the most significant driver for a segment-specific data model is the ease of monitoring changes. Once the road network has been divided into road segments of equal length, which are used to store data, monitoring changes will in practice be simply a matter of making subtractions from data collected during different runs. As the boundaries of each road segment are precisely defined with regard to their location, the method will still encounter the following challenge in close proximity to segment boundaries: If the damage is on the boundary of two road segments and the accuracy of the location data is out by even ten centimetres, this could cause the hole in question to be assigned to two different segments during different runs. Neither can this allocation problem be eliminated by using more accurate locationing or increasing the resolution of the road segment, even though the latter would reduce the number of segment boundaries in relation to the entire network.

A window function could be one solution to the road segment allocation problem with regard to monitoring changes. Instead of comparing road segments directly to each other, neighbouring segments should also be taken into consideration, that is, the road segments immediately before and after. A window function, combined with the correct choice of resolution, would enable a solution that could utilise cost-effective consumer-level data collection equipment. Each segment would also contain a maintenance class that would specify the subsequent coverage requirements and summary reporting publication cycles.

The following exceptions must also be taken into consideration in order to create reliable summary reports.

- The coverage of the footage must exceed a certain road-specific limit value in order for that road to be included in the report. Otherwise, results from the previous monitoring period should be used.
- If a road is so full of damage that it is impossible to repair, it should be separated from the rest of the network for the purposes of monitoring, so that such roads do not end up dominating the monitoring metrics.

7 Further development

7.1 Developing the procurement model

When there is an available observation tool to describe the amount of damage or the condition of the road, pavement repair can be procured, either wholly or in part, via target-price or total-price service contracts rather than the current system of unit-price contracts. The condition of the road network, measured according to the amount of damage, would then act as the primary functionality requirement. This would require a way of describing the condition of the road appropriately and with sufficient accuracy. As the main repair procedure is patching a hole in the road surface, the design of the contract procurement model should pay particular attention to the prevention of holes and the effective repair of any holes that do form before the damage poses a danger to road safety. The procurement model must steer contractors' operations in a way that it is worthwhile for contractors to prevent the formation of holes and repair any holes in good weather outside the snow season. The model should also encourage development and continual monitoring of the status of the road network, which will enable more efficient operations.

Before a new procurement model can be developed and introduced, criteria will be required to assess the procurement model's functionality and effectiveness compared to existing unit-price models. Example criteria could be:
- total cost
- distribution of repair work over the seasons
- distribution of manual and machine repairs
- amount of Harja feedback and number of claims for damages
- checking the repeated repair of the same instance of damage.

7.2 Developing the data model

This study outlined the basic logic of the data model and justified its usefulness. The practical implementation of the data model and its application to reporting will now continue as part of Vaisala RoadAI's development work. The application of the data model to monitoring pavement repair contracts in the region administered by Pirkanmaa ELY Centre will remain an area for further development.

7.3 Auditing the performance of computer vision

In the study conducted during the pilot, the identification and classification abilities of humans and computer vision were close to each other. The functionality of the technology cannot be disputed. However, from an overall perspective, it is important that the quality of computer vision can be demonstrated and audited in a competitive situation. Rather than handling computer vision as a separate method, auditing should be based on a method that can also be performed manually by a human. This would raise the level of abstraction and auditing would focus on comparing the final results.

Additionally, the quality assurance process would not be overly tied to a particular provider's technical solution. A manual solution performed by a human would work as an auditing regulation method and also ensure that the method remains sufficiently straightforward. In addition, it would be beneficial to develop the process in a direction that would enable new technology yet not make it necessary.

7.4 Developing tools

An observation tool and the ability to verify damage are both essential features of the user interface, but are by no means the only use cases in a pavement repair contract. A dashboard view that displays both current and history data is the most important summary for the parties involved in a contract. An additional requirement was discovered during the pilot: the user interface should enable users to mark repairs that have been made and correct any mistakes in the computer vision's interpretations.

Another key observation that arose in feedback obtained from contractors during the pilot was that it would be handy if the mobile app could be developed in a direction that would enable entries about repairs to be made whilst out in the field.

7.5 Specifying and monitoring hole identification

The ability to uniquely identify holes from video footage requires the reliable tracking of holes between frames. The number of identifications poses a challenge: For example, whilst road signs can already be seen from a distance and are identified in numerous consecutive frames, a hole is often only identified in 1–2 frames taken very close to the hole. The main reason for this is that holes form on the road's surface, while road signs are designed to be placed at a ninety-degree angle to the road's surface, so that they can be seen from as far away as possible.

Uniquely identifying holes and a segment-specific data model are not mutually exclusive options. Instance-level (individual) monitoring would increase the accuracy of a segment-specific data model, thereby enabling information about the number of holes to be added to the segment.

7.6 Integration into other systems

The system used in the pilot currently enables data to be exported as a csv/excel file, for the purposes of watching the road videos and verifying the accuracy of the data displayed via the user interface. However, at some point it will be necessary for the data's final storage place to be in a national system, such as Harja.

8 Conclusions and summary

Developing a new procurement model for pavement repair contracts will require a reliable method for measuring damage or the condition of the road. The planned procurement model aims to promote the development and introduction of new repair and quality assurance technology. The objective is a method that would encourage contractors to engage in efficient and correctly timed repair work of a high standard.

Transferring to a new procurement model will require a way of reliably indicating both the surface condition of the road network and any changes in its condition. Computer vision offers excellent potential in this area, and would be cost-effective as well as easy to mobilise. From the perspective of the data production process, the data generated by Posti's vehicles is best suited to fulfilling the coverage requirements of roads in maintenance category Y1, as the data should be continually updated and the service level requirement is high. Footage to monitor changes in the condition of roads in maintenance categories Y2 and Y3 would be best filmed using either the contractor's or another service provider's vehicles.

The pilot sought to determine whether artificial intelligence could be used to continually monitor the condition of paved roads. On the basis of the results, computer vision performed at a good level in identifying holes and cracks. The system used in the pilot is already generating added value thanks to its up-to-date road videos and the accuracy of its observations, even though it is not yet suitable for full point-specific monitoring.

Computer vision was compared with observations made by three humans. When it came to situations in which a human identified a hole but computer vision did not, the average of the three humans' identifications was only 2.3 per cent for severe holes and 10 per cent for moderate holes. Greater variation in accuracy was seen when defining the severity of damage. Defining whether the damage was severe or moderate was clearly a challenging task for the humans, and thereby was not a straightforward task for the computer vision to learn, either. In light of the results, comparison between the humans or with the computer vision generated similar variation when defining the severity of the damage.

In order to ensure the development of a high-quality method, it would be good to continue the use of technology in pavement repair contracts and to create processes to harmonise the differences in interpretation between humans and computer vision. When it comes to reporting, the pilot indicated that a data model based on road segments is important for monitoring hole formation in roads. The data model is based on sections of road 5–10 metres in length, which provide sufficient coverage for data production and enable the condition of the road surface to be monitored. They also enable a more accurate reporting method to be used for summary reporting.

During the pilot, it became clear that the video footage produced for quality assurance in pavement repair contracts was also suitable for use – and worthwhile using – in scheduling pavings. Up-to-date footage is extremely useful in planning paving work, and it enables the final prioritisation and assessment of paving contracts to be carried out in a straightforward manner.

In addition, it has clear and significant financial added value in general, as the road videos are produced in conjunction with existing operations.

Computer vision worked sufficiently well with respect to its practical application. Although the results were not perfect in comparison to humans, it is justifiable to say that further research will only be able to achieve minor improvements in harmonising the results. In order to genuinely start seeking the desired level of cost-effectiveness, it would be worthwhile to switch from research and pilots to operative use. With respect to developing a procurement model and practical processes for pavement repair contracts, the optimal approach would be to progress gradually towards the target state. In the new model, computer vision would function as a continuous source of information about the condition of the road, producing data at specified intervals. A human validation process would run alongside this for as long as necessary for adequate machine learning to occur. This kind of master-apprentice model will help to achieve a genuinely objective and high-quality 'app robot' that can independently supervise the condition of the paved road network.