Pothole detection in the woods: a deep learning approach for forest road surface monitoring with dashcams

Mostafa Hoseini, Stefano Puliti, Stephan Hoffmann & Rasmus Astrup

Published online: 18 Dec 2023.

Submit your article to this journal ↗

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Pothole detection in the woods: a deep learning approach for forest road surface monitoring with dashcams

Mostafa Hoseini (ID), Stefano Puliti (ID), Stephan Hoffmann (ID), and Rasmus Astrup (ID)

Division of Forest and Forest Resources, Department of Forest Operations and Digitalization, Norwegian Institute for Bioeconomy Research, Ås, Norway

ABSTRACT

Sustainable forest management systems require operational measures to preserve the functional design of forest roads. Frequent road data collection and analysis are essential to support target-oriented and efficient maintenance planning and operations. This study demonstrates an automated solution for monitoring forest road surface deterioration using consumer-grade optical sensors. A YOLOv5 model with StrongSORT tracking was adapted and trained to detect and track potholes in the videos captured by vehicle-mounted cameras. For model training, datasets recorded in diverse geographical regions under different weather conditions were used. The model shows a detection and tracking performance of up to a precision and recall level of 0.79 and 0.58, respectively, with 0.70 mean average precision at an intersection over union (IoU) of at least 0.5. We applied the trained model to a forest road in southern Norway, recorded with a Global Navigation Satellite System (GNSS)–fitted dashcam. GNSS-delivered geographical coordinates at 10 Hz rate were used to geolocate the detected potholes. The geolocation performance over this example road stretch of 1 km exhibited a root mean square deviation of about 9.7 m compared to OpenStreetMap. Finally, an exemple road deterioration map was compiled, which can be used for scheduling road maintenance operations.

Introduction

Forest roads are a central part of the forest infrastructure that allows for access and management of the forest (Boston 2016). These roads are commonly designed as haul roads to relocate machinery and truck timber during the harvesting phase, but also support other management activities such as stand establishment and tending, monitoring, or surveying tasks that require access for personnel and equipment. From a societal perspective, functional forest roads are important for recreational purposes, but also grant access during emergencies such as rescue or firefighting (Waga et al. 2020). To safeguard the unrestricted functionality of a forest road according to its category (e.g. permanent haul road, granting all-year access with logger trucks), periodic and preventive maintenance is essential to ensure management of a sustainable forest road network (Ezzati et al. 2021).

Maintenance requirements in terms of intervals and magnitude can differ significantly according to road standard and building material, traffic intensity, and e.g. occurring extreme weather events, with climate change–induced differing in frequency and patterns (Dodson 2021). Road maintenance costs are considerable, but negligence in maintenance eventually leads to consequences such as road failure or restricting a road's functionality and increasing the risk of negative environmental impacts (Dietz et al. [1984] 2011; Erber et al. 2021). Consequently, forest roads need to be frequently monitored to detect wear and maintenance status. Surface deteriorations in

the form of potholes are often the most obvious signs of road wear since these also affect driving comfort and travel speed. If not attended to in time, potholes can lead to greater defects and are also a good indicator of lack of crossfall and other road drainage issues (New Zealand forest road engineering manual 2020). Potholes retain water, which can then penetrate the aggregates, causing strength loss of the road base layer. A common consequence of such seep from potholes is bearing capacity loss and frost heave, which can cause complete road failure (Saarenketo and Aho 2005). To avoid cost-intensive road repairs, comprehensive and frequently updated information about the road surface to determine the kind and urgency of required actions is essential for a sustainable forest road management system.

Forest road monitoring is currently widely practiced by manual inspection, which is time consuming and expensive, and the results are often subjective and dependent on the surveyor (Susnjar et al. 2020). Attempts to formalize these procedures through specified assessment protocols and derived indices, such as the Forest Road Pavement Condition Index (Heidari et al. 2022), did not gain wide acceptance, mainly due to the continuing high labor input. Consequently, maintenance operations are often scheduled in an inefficient manner. Technological progress in sensing technologies has made its way to forestry, and offers numerous potential ways to efficiently gather and process road data to be incorporated in decision support systems (Talbot et al. 2017; Picchio et al. 2019). Implementing sensor technologies into operational traffic of,

for example, timber trucks, offers time and cost-efficient alternatives to manual road surveys and support target–oriented maintenance operations.

Among highway authorities, dedicated survey systems of multiple sensors and high resolution have already been established to detect various wear-induced road anomalies of concern to traffic safety and road durability (van der Horst et al. 2019). Such approaches can be integrated into mobile platforms, such as dedicated surveying vehicles of commercial operators (e.g. Roadscanners Oy, Finland), or even can be conducted with stationary build-in sensors to create "smart roads" (Barriera et al. 2020). Yet, forest roads are low-volume roads, requiring less cost-intensive systems, with lower resolution often being sufficient. Early attempts to use a rather simple sensor platform, combined with a maintenance planning tool, date back to the early 2000s with the Canadian Opti-Grade system making use of acceleration sensors to determine road deterioration based on surface roughness (Brown et al. 2003). With acceleration and Global Navigation Satellite System (GNSS) sensors available on modern smartphones, this approach has been widely transferred to monitor rural roads through end-user devices with tailored applications (e.g. roadroid.com), including first trials of transferring these techniques into a forestry setting (Susnjar et al. 2020).

Although smartphones are economic, bias due to vehicle calibration, driving habits, and phone positioning, as well as general limitations of accelerator sensors, allow a road condition assessment at only very coarse resolution (Sattar et al. 2018). For instance, it is not possible to determine if the vibrations originate from a pothole or are caused by uprising material from the subgrade such as larger stones. But this is relevant information for target-oriented maintenance planning. On the other hand, high resolution data as, for example, that captured with mobile laser scanner systems used among highway authorities, are too costly and sophisticated for low traffic volume roads (Cao et al. 2020). Despite some first promising pilot approaches to downsize laser scanning techniques to forestry road applications (Ferenčík et al. 2019), this remains a technology associated with higher efforts. Thus, optical systems and image-based object detection and recognition algorithms can be a suitable alternative for forestry purposes, with sufficient information provision at moderate levels.

Maintenance monitoring for potholes on paved public roads with camera solutions has received considerable attention (Roberts et al. 2020; Cao et al. 2020) since it is seen as a good compromise between cost-efficiency and data resolution compared to other sensors (Kim et al. 2022). The ability to integrate such solutions into smartphones allows a comprehensive and extensive road survey through daily traffic by, for example, postal services or public vehicles, constantly updating and increasing databases, as is already widely implemented among Japanese municipalities (Maeda et al. 2018; Arya et al. 2021).

Bounding box detectors such as the popular you-only-look-once (YOLO: Redmon et al. 2016) are often the essential base of such monitoring systems, and are also increasingly used within forestry computer-vision applications (Puliti et al. 2022; Puliti and Astrup 2022). The use of convolutional neural networks (CNNs), or deep learning, is motivated by the high quality of the output especially in computer vision tasks and by the ease of deployment in real-world scenarios thanks to the real-time inference capabilities. In addition, repositories, such as Ultralytics' YOLOv5 (Jocher et al. 2022), provide technological access, allowing a broad community to train and deploy custom models efficiently while obtaining state-of-the-art results.

The range of YOLOv5 applications covers diverse domains, such as autonomous vehicles (Benjumea et al. 2021), surveillance systems (Ali et al. 2022), and medical imaging (Mohiyuddin et al. 2022), showcasing its versatility in addressing various computer vision challenges. YOLOv5 as an open-source, fast, and efficient one-stage detector has also made its way into forest road applications. A first attempt for an automized visual inspection system for forest roads was done by Starke and Geiger (2022), who applied a YOLOv5 computer vision-based model to detect forest road surface deterioration with consumer-grade optical sensors. However, this low-resolution approach based on smartphone and tablet images was able to detect only waterlogged sections on the road. Heidari et al. (2022) also presented a method for identifying the pavement damage through the processing of smartphone images using the YOLO algorithm.

Bounding box detectors are useful for frame-by-frame detection; however, when dealing with videos one must consider the sequential order of the frames and the need to associate the detected bounding boxes to separate instances across frames. Object tracking techniques allow for tagging the bounding box instances across frames and, consequently, counting the number of unique objects and mapping them in space. Tracking-by-detection methods have traditionally been a popular choice (Bochinski et al. 2017) in the domain of online multiobject tracking (MOT); that is, algorithms allowing the real-time tracking of multiple instances. These tracking methods first leverage an object tracker to locate objects in single frames followed by the association of these detected objects across frames using geometrical, motion, and appearance information. Recent implementations of the popular Simple Online and Realtime Tracking method (SORT; Bewley et al. 2016), have been readapted to complement the objects' geometrical information (i.e. size and dimensions of the objects) with an appearance model, a deep learning model useful for reidentification of instances across frames (StrongSORT; Du et al. 2022).

In this study, we harnessed the power of the introduced deep learning–based object detection and tracking tools to process the video recordings of a consumer-grade vehicle dashcam for detecting and geolocating potholes as indicators of forest road surface deterioration. Moreover, our developed approach endeavors to generate a deterioration map, which can serve as a valuable resource for effectively scheduling road maintenance operations. In addition, the presented method can be used to generate input variables for modeling road deterioration processes as, for example, conducted through logistic regression analysis and artificial neural networks (Heidari et al. 2018).

Materials and methods

Although not geographically restricted in its application, the developed road-monitoring approach focused on Norwegian conditions for standard secondary low-volume forest roads of class 3, according to the Norwegian standard for agricultural roads (Lanbruksdirektorat, 2016). These roads follow a geometry to permit safe and efficient driving of haul trucks and must be able to be trafficked with loads year-round with limitations only during periods of heavy rainfall or freeze and thaw cycles in spring.

Video data

Road footage used for this study primarily relied on videos captured on various forest roads in southern Norway (Viken county), using the popular GNSS-equipped 622GW dashcam (NEXTBASE, UK), hereby referred to as "dashcam data" (see Table 1 and Figure 1). These videos were recorded with the camera positioned in the upper center of the windshield, capturing the oncoming road while driving at a moderate speed with a focus on 20 km/h. For recording, the camera was set at a resolution of 1080P (1920×1080 pixels) and a recording speed of 30 frames per second (fps), with engaged image stabilization and a polarization filter on the lens. The videos were captured under varying weather conditions and during different times of the day with the aim of covering a broad range of real-world scenarios.

To increase the variability and size of the dataset, we augmented the dashcam data with video frames captured for a different purpose on Norwegian forest roads in the same region using a tablet, hereby referred to as "tablet data," as well as an already annotated dataset from South Africa in an urban setting (Nienaber et al. 2015, 2015).

Data annotation

The video frames from Norwegian forest roads captured with the dashcam and tablet (Table 1) were manually annotated by different teams of annotators. The annotation consisted of drawing bounding boxes around each of the visible potholes.

The annotations were mostly performed at 1 fps to increase the number of videos that could be annotated within the given time frame.

For the test data (Table 1), the annotation also included the instance unique identifier allowing us to track the potholes across sequential frames and, consequently, to evaluate the object tracking algorithm. To allow for smooth tracking of the potholes, the test videos were annotated with a higher sampling rate (15–30 fps) compared to the modeling dashcam data and, thus, resulted in a substantially larger number of annotated potholes. Figure 2 demonstrates some of the annotated potholes.

Test data

The test data consisted of two videos captured over the same road stretch (approx. 300 m) on a lowland timber haul road in southern Norway (Våler Kommune/Viken county), in a typical spruce (*Picea abies*)–dominated Nordic coniferous forest. Such a road stretch was selected as it was representative of various forest stand conditions from dense to partly open, and covered different appearances of potholes from being shallow, in their early development, to deep and established, already affecting the base layer (Figure 2). In addition, to account for various weather conditions, the road was filmed twice. The first recording was conducted on a bright and sunny day, and the video was characterized by solar reflections on the windshield and dried out potholes. The second recording took place after a rainy night and was characterized by well-visible and water-filled potholes, as well as by low-light conditions due to overcast sky.

Dashcam GNSS data

The deployed dashcam allowed for integrating geolocation information to the video recordings. The geolocation information was received at a fixed rate of 10 Hz and was in the standard format of NMEA-0183 (National Marine Electronics Association). The main NMEA message used in the dashcam for delivering the geolocation information was the "GPGGA"

Table 1. Summary of the different data sources used to train the pothole detector with corresponding number of images and potholes.

Dataset name	Modeling		Test	
	Images	Potholes	Images	Potholes
Dashcam	2779	6676	1726	3788
Tablet	376	561		
Nienaber	3888	9428		
Total	**7043**	**16,665**	**1726**	**3788**

Dashcam   Tablet   Nienaber et al. 2015

Figure 1. Examples of different data sources used for model training. The dashcam data was further split into validation and test sets.
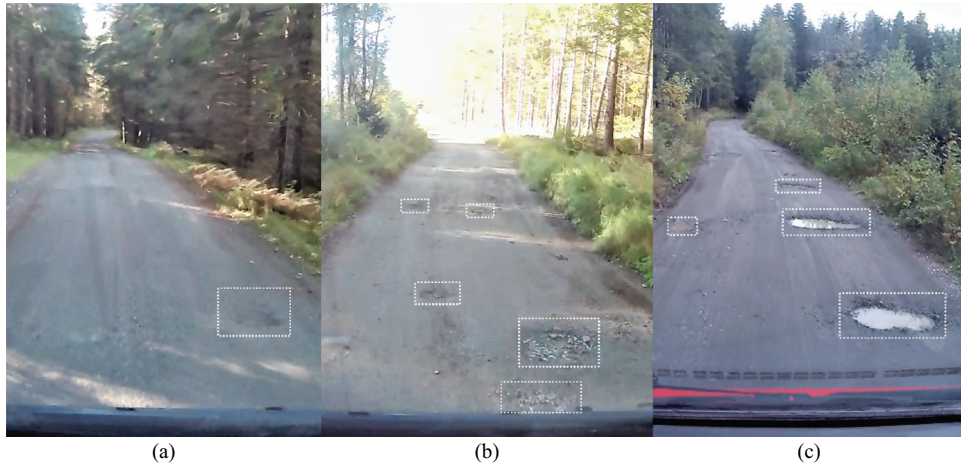
Figure 2. Exemplary pothole conditions represented on the test road: (a) light roadway depression (dry conditions) in the wearing course, (b) deep pothole with dispersed aggregates of the base layer, and (c) water-filled potholes.

message known as Global Positioning System (GPS) fix data. The message included different pieces of information from which we extracted and used the measurement time in UTC, latitude, longitude, number of satellites used for the calculation of the position, and horizontal dilution of precision (HDOP) as an indicator of the predicted overall quality of the position in the horizontal domain. Since in the forest environment the dashcam's GNSS receiver can frequently lose the satellite signals and deliver erroneous positions, a data screening was required to exclude any possible outliers from the processing.

Methodological pipeline

The overall AI pipeline developed by this study (Figure 3) involves different processes that can be grouped into the following main steps:

(1) **Pothole detector training**: a YOLOv5 pothole detector was trained using our modeling data.
(2) **Pothole tracking**: the trained detector was used to detect potholes in the test videos and the detected

potholes were tracked through frames and aggregated into single instances.
(3) **Pothole mapping**: the tracked potholes were assigned geographic coordinates and the potholes were dinned into 20 -m road sections.

Pothole detector training and validation

A YOLOv5 model was trained using the annotated modeling data. For the training, 10% of the annotated frames from the modeling dashcam data was used for validation and selection of the best model. Given that this currently represents a server solution (i.e. not for edge computing), we opted for a large model (i.e. YOLOv5x) with an image size of 1280, thus maximizing the performance.

We trained the model on a Linux machine with a single Tesla V100 with 8GB of memory. Based on these specifications, the largest batch size was two.

The resulting detector was validated against the validation and the test dataset using common object detection metrics such as the precision (P), recall (R), and mean average precision at intersection over union (IoU) of 0.5 (mAP@0.5). To
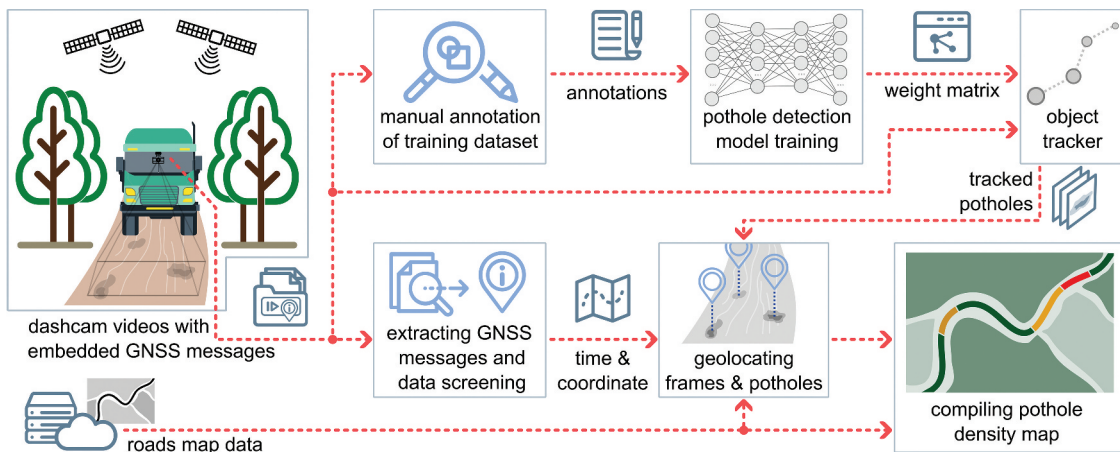

Figure 3. The developed pipeline for monitoring forest road surface deterioration using a GNSS-augmented dashcam. The depicted processing steps were carried out using dedicated python scripts developed in this study. The flowchart was created in inkscape software.

Table 2. The standard file format for multiobject tracking (MOT) methods' output.

Column	Header title	Description
1	frame	The frame number (1-based)
2	id	A unique identification number which is assigned for each object instance (1-based)
3	bb_left	Left edge coordinate of the object's bounding box (pixels)
4	bb_top	Top edge coordinate of the object's bounding box (pixels)
5	bb_width	The width of the object's bounding box (pixels)
6	bb_height	The height of the object's bounding box (pixels)
7	conf	Detection confidence (ranging from 0 to 1)
8	x	x-coordinate (ignored for the 2D case)
9	y	y-coordinate (ignored for the 2D case)
10	z	z-coordinate (ignored for the 2D case)

assess the model's performance under varying atmospheric conditions present during the capture of the test videos, we evaluated the two videos separately in addition to reporting the overall test data performance.

Pothole tracker implementation and validation

In this study we utilized the StrongSORT OSNet (Broström 2022) tracker, given its direct integration with YOLOv5 models. The repository includes a main function to run several tracking algorithms given an input video and model weights for a pretrained YOLOv5 model. The output of the tracking consists of a MOT compliant text file (MOT Challenge 2023) containing the columns listed in Table 2.

We opted for the default StrongSORT (Du et al. 2022) with the OSNet (Zhou et al. 2019) as the reidentification model (Bochinski et al. 2017). The performance of the object tracking can be summarized by Higher Order Tracking Accuracy (HOTA) metrics (Luiten et al. 2021). HOTA is proposed to address difficulties of evaluating multiobject tracking (MOT) methods by unifying three metrics that measure the detection, localization, and association performances based on IoU. The IoU score, also known as the Jaccard index (J), varies between zero and one and is calculated as the intersection of the detection and annotation areas divided by the union of the two areas (Eq. 1):

$$IoU_{A,D} = J(A, D) = \frac{|A \cap D|}{|A \cup D|} \quad (1)$$

where A and D, respectively, refer to the areas of annotated and detected objects. The overall success rate of a method regarding the spatial alignments of prediction against ground truths are captured by localization accuracy (φ) (Eq. 2), which is the average IoU score of all successfully detected potholes:

$$\varphi = \frac{1}{N_{TP}} \sum_{k=1}^{N_{TP}} J_k \quad (2)$$

with N_{TP} being the number of true positive (TP) cases (i.e. the successful detections) identified with a minimum IoU of 50% (i.e. $\varphi_0 = 0.5$). Such a detection is then considered as a common object in the detection and ground truth sets. Consequently, the detection accuracy (α_D) (Eq. 3) is the ratio of successful detections counted to the total number of objects in the union of the two sets:

$$\alpha_D(\varphi_0) = \frac{N_{TP}}{N_{TP} + N_{FN} + N_{FP}} \quad \varphi \geq \varphi_0 \quad (3)$$

where FN indicates false negative (i.e. when an existing pothole is not detected by the model), and FP indicates false positive, which refers to the case when the model mistakenly reports a detection. Finally, association accuracy for an object tracker determines how well the tracker links detections over time into the same identities (IDs). The overall association accuracy (α_A) over all successful detections (N_{TP}) is defined as (Eq. 4):

$$\alpha_A(\varphi_0) = \frac{1}{N_{TP}} \sum_{i=1}^{N_{TP}} \frac{N_{TPA}}{N_{TPA} + N_{FNA} + N_{FPA}} \quad \varphi \geq \varphi_0 \quad (4)$$

with the letter A in the subscripts referring to the association. The readers are referred to Luiten et al. (2021) for a more detailed description of the HOTA metrics.

Geolocating potholes

The dashcam used in this study has a built-in GNSS module that can track satellite signals from two different GNSS constellations. Although this feature significantly increases the number of observations used for the coordinate calculation, the forest environment can still leave its degrading signature on the position accuracy. Therefore, an outlier detection method was applied to the geolocation information that was extracted from the dashcam videos. In the first step, gross errors (i.e. coordinates or time stamps falling outside the spatiotemporal scope of the validation dataset) were eliminated. Such errors have been found in the GNSS module's output mainly during periods when the number of satellites has significantly dropped. In the second step, a sliding window of 10 observations accounting for 1 second worth of observations was applied to latitude, longitude, and time values to calculate moving average and standard deviation of each parameter. If the difference of the variable with respect to the corresponding moving average exceeded the threshold of three standard deviations, the coordinates and time of that epoch were flagged as an outlier and were excluded from the processing.

Besides gross errors and outliers, the Standard Positioning Service (SPS) of GNSS that is generally implemented and used by low-cost receivers can have a significant positioning error. Therefore, the dashcam trajectory coordinates were orthogonally projected to the nearest forest road retrieved from OSM (OpenStreetMap 2023) data for the mapping purpose. The projected coordinates (at a fixed rate of 10 Hz) were then used in a linear interpolation process to calculate the

Table 3. Exemplary classification of forest road surface deterioration based on pothole density.

Class	Potholes density	Operational need	Color code
Category I	Low	The road requires no maintenance work	Green
Category II	Moderate	The road should be considered for preventive, ongoing maintenance work (e.g. grading of road surface)	Yellow
Category III	High	The road is due for periodic maintenance (e.g. resurfacing)	Orange
Category IV	Very high	The road is in urgent need of resurfacing and limitations for the design traffic needs to be currently accounted for	Red

geolocation of video frames at the rate of 30 fps. Each pothole instance was usually tracked within several frames and was tagged with a unique identifier. Therefore, the nearest available coordinates (i.e. coordinates of the latest frame in which the pothole appeared) were used as the pothole's geolocation.

Classification of road deterioration

For road maintenance purposes, a favorable representation of the georeferenced potholes can be derived in the form of a pothole density map. To this end, along-road distance of each pothole instance with respect to a reference point on the road was calculated. Then, the potholes were grouped and counted based on a road segmentation scheme with a predetermined length. From a road manager's perspective, it is of interest to classify road units into urgency of required attendance to ensure the maintenance of their functional class by scheduling required work accordingly. Thus, we classified the road units into four categories as represented in Table 3.

In this study, we used a segmentation length of 20 m and classified segments with up to two potholes as Category I, three to five potholes as Category II, six to nine potholes as Category III, and 10 potholes or more as Category IV. This classification is limited to the pothole count and does not fully picture the overall road condition, nor does it consider the different magnitudes of potholes. It also does not follow any official

categorization of forest road maintenance standards, but it is a first approach to evaluate the road condition. We defined these categories purely to convey the concept of a maintenance classification system based on a pothole count, which allows high flexibility to adapt according to different road classes and local standards by decreasing or increasing the pothole number thresholds.

Results

Pothole detector validation

The training of the pothole detector was interrupted after 299 epochs, following 100 epochs without model improvement (see Figure 4). Based on the internal validation data, the best model was found for epoch 199 corresponding to a validation precision (P), recall (R), and mean average precision at IoU of 0.5 (mAP@0.5) of 0.62, 0.53, and 0.55, respectively.

Pothole tracking validation

An evaluation of pothole tracking performance using HOTA metrics is reported in Table 4. The HOTA scores confirm a significantly different detection performance for the two validation cases. For the case of wet and overcast weather with water-filled potholes, the detection was approximately
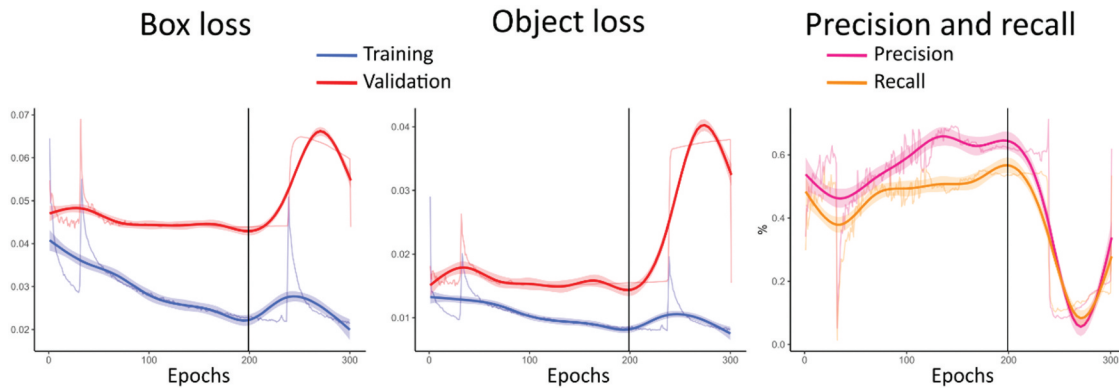

Figure 4. Summary of the training metrics development over the epochs. The black vertical line represents the 199th epoch. The validation data correspond to the subset of the modeling data.

Table 4. Performance assessment report for the pothole detection approach used in this study based on the Higher Order Tracking Accuracy (HOTA) metrics as a tool for evaluating multiobject tracking (MOT) models. All the entries are in percentages.

Condition	HOTA	Localization accuracy	Detection accuracy	Detection recall	Detection precision	Association accuracy	Association recall	Association precision
Sunny day	18.60	71.21	13.51	17.82	31.06	25.75	46.30	32.57
After rain	31.39	77.23	34.33	41.54	56.79	29.02	55.86	34.54
Combined	29.76	76.77	30.98	38.03	53.71	28.92	55.48	34.57

Table 5. Number of pothole instances identified by model prediction and manual annotation under two weather conditions. On the sunny day, the model correctly detected 30.2% (13 out of 43) of the ground truth potholes compared to 43.2% (35 out of 81) on the overcast day.

Sunny day		Model prediction (19 detections)		Overcast (after rain)		Model prediction (44 detections)	
		Pothole (Positive)	Background (Negative)			Pothole (Positive)	Background (Negative)
Reference (43 potholes)	Pothole (Positive)	13	30	Reference (81 potholes)	Pothole (Positive)	35	46
	Background (Negative)	6			Background (Negative)	9	

boosted by a factor of ~ 2.5 compared to the video recorded on a sunny day. In contrast, the metrics reporting on the association performance did not highlight any prominent supremacy of the wet and overcast weather condition, suggesting that the association algorithm is not sensitive to the weather difference. However, the overall HOTA score was clearly affected by the different detection efficiency of the two cases.

Table 5 summarizes an important feature for overall performance assessment, which is the number of detected and mapped pothole instances compared to the ground truths (i.e. the manually annotated potholes). The table highlights that an overcast weather condition after some rainfall can significantly boost the performance of the method.

Pothole mapping performance

An assessment of the quality of geolocation information delivered by the dashcam's GNSS module is summarized in Figure 5. Panel (a) of the figure shows the overlaid graphs of the original coordinates received from the GNSS module, and the projected coordinates to the corresponding forest road. The trajectory shows a good agreement with the road; however, there is a varying deviation from the road that is depicted in Figure 5(b). The deviation over the illustrated road stretch can reach to about 33 m with a root mean square (RMS) value of about 13.7 m over the road stretch, combining the uncertainties of the GNSS coordinates and the OSM road data. It should be noted that the OSM data over the selected road stretch show an RMS deviation of 1.7 m with respect to the coordinates from a high-precision geodetic GNSS receiver in another experiment. The parameters plotted on Figure 5(c) can predict the expected overall quality of the delivered coordinates.

In Figure 6, the plots in left panel display the number of potholes per video frame, with the horizontal axis indicating the road length relative to the first frame. Meanwhile, the right panels in Figure 6 depict road deterioration maps based on the four categories described in Table 3. These plots were generated using manually annotated potholes (ground truth data), and the detection results for two test days: a sunny day and an overcast day following rainfall.

Discussion

The analysis of the training results revealed that the F1-score was maximized at a confidence value of 0.232 and, thus, such a value was used to threshold predictions in the test data. When validated against the test data, the model showed an increase in all metrics with values for P, R, and mAP@0.5 of 0.75, 0.57, and 0.68, respectively. Such a performance boost was mainly driven by the model performance under overcast and wet conditions (P = 0.79; R = 0.58; mAP@0.5 = 0.70). In contrast, the performance during sunny conditions (P = 0.60; R = 0.45; mAP@0.5 = 0.47) was poorer than what was found against the validation data. Such contrasting performances between the sunny and overcast conditions agrees with the manual image annotation experience. The experience confirms that the visibility of potholes in the footage was more obvious following rainfall, when the potholes were filled with water, and lower reflections from the windshield were experienced due to the overcast sky.

Reflecting on the resulting road deterioration map, although being similar in pattern with road sections being more worn out than others, according to the exemplary road damage classification, the overall result was different. Whereas during the sunny conditions the road was in overall good shape, with only one "hot spot" section in Category IV, and none in Category III. For road management, such a result would require just a further monitoring
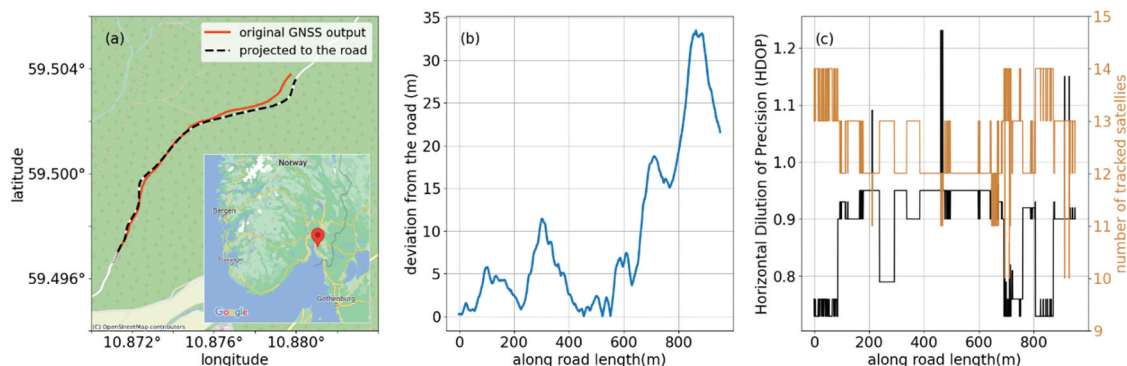

Figure 5. (a) The original GNSS coordinates (after removing outliers) overlaid by projected points to the road, (b) the deviation of original coordinates from the projected points, and (c) horizontal dilution of precision (HDOP) and satellite count graph on separate y-axes.
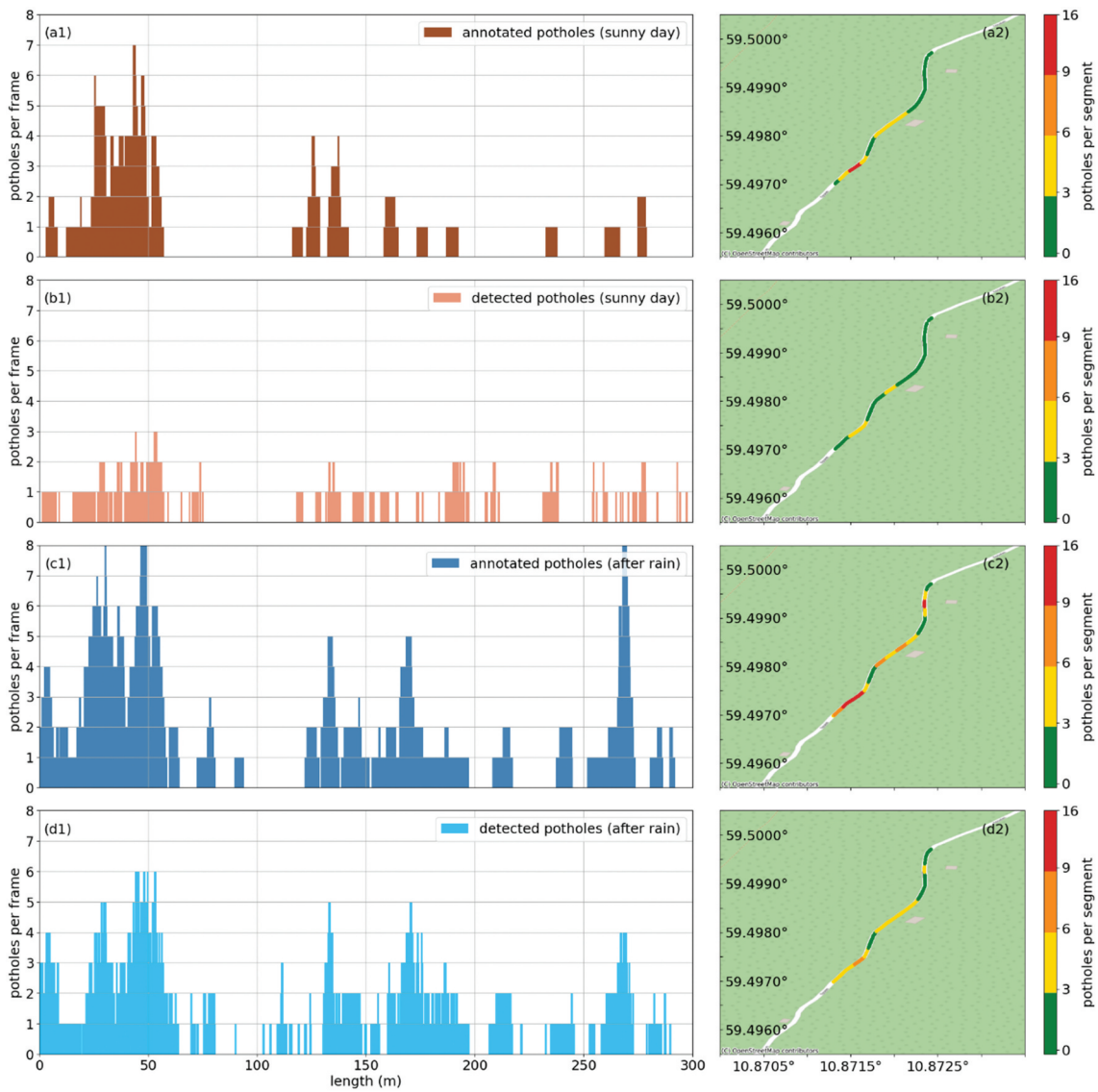
Figure 6. Panels on the left: the pothole detection results presented in terms of the number of detections per video frame translated to relative length of the road. A pothole instance can be detected and tracked in several consecutive frames. Panels on the right: road deterioration map produced based on the classification of number of pothole instances per 20-m segment. Panels (a1)-(a2) and (b1)-(b2) are associated with the dashcam video captured on a sunny day, and the results shown on (c1)-(c2) and (d1)-(d2) are based on a video recorded during an overcast weather condition after rain.

throughout the operational season, or at most some spot improvement work. In contrast, assessing the same road after a rainy day reveals a different situation, i.e., a maintenance status which already limits the technical functionality of the road. The detection of extended sections in maintenance Categories III and IV, classified the road due for periodic maintenance, and work such as resurfacing needed to be scheduled soon.

These two contradicting outcomes from a road manager's perspective could lead to false planning of resources and operation scheduling. Thus, limitations of optical sensors referred to sunlight reflections and poor visibility of light surface depressions under the canopy, when potholes are not water filled, must be accounted for. Consequently, the users of such an application need to factor in the conditions under which they can retrieve the best results. Furthermore, it is important to see our study as a first step toward the operational application, where different than in our study, a single road stretch may be driven through several times, thus providing the opportunity to average multiple

tracking results over time and thus reducing the uncertainty of the pothole maps.

The quality of the final maps can also be influenced by another parameter – the positioning of the potholes. Accurately geolocating detected potholes depends heavily on the performance of the GNSS unit in the dashcam. The GNSS metadata highlight a high number of tracked satellites with HDOP values mainly below 1, suggesting an overall favorable geometry for calculating the position. Nevertheless, the number of satellite signals that can uninterruptedly reach to the modest antenna of the GNSS module may not be as high as what is shown in Figure 5(c). Our visual inspection confirmed that the local minima of the deviation graph in Figure 5(b), as an indicator of GNSS coordinates error, well corresponded to more open sky views in the forest.

Future work is planned to create a seamless pipeline that starts from transferring the captured videos to a cloud infrastructure on a near real-time basis. The videos uploaded to the cloud are instantly processed to update the current road

deterioration maps. Availability of such a pipeline for the forest roads with active users can provide a frequently updated observation source for constantly improving the detection model and the accuracy of the generated deterioration maps.

Conclusions

An automated approach for monitoring forest road surface deterioration in terms of potholes was demonstrated in this article. The approach leverages computer vision to develop a model for detecting potholes in the captured videos of a standard dashcam. The videos have embedded geolocation information provided by a dual-constellation GNSS module. The results of applying the developed model to the video recordings under different weather conditions highlight an overall promising performance. However, a significantly higher success rate was observed for the case of combined overcast and wet weather, which created water-filled potholes. Further conditional factors, such as degree of pothole development and how these are affected by the various lighting conditions, should be considered for future studies. An exemplary road deterioration map (limited to pothole features) was demonstrated that can be frequently updated and utilized for preventative maintenance purposes.

The study also revealed some potential improvement opportunities. In case of a need for higher positioning accuracy for the potholes, a combination of hardware and processing enhancements can be considered for future studies. A dashcam with an external GNSS antenna input could improve signal reception and reduce the number of outliers in the GNSS-delivered coordinates. Moreover, a monocular depth estimation method could be assimilated to the processing for transferring the GNSS receiver's location to the detected potholes.

ORCID

Mostafa Hoseini http://orcid.org/0000-0002-1284-3422
Stefano Puliti http://orcid.org/0000-0003-4624-8987
Stephan Hoffmann http://orcid.org/0000-0002-8077-967X
Rasmus Astrup http://orcid.org/0000-0003-2988-9520

References

Ali L, Alnajjar F, Parambil MMA, Younes MI, Abdelhalim ZI, Aljassmi H. 2022. Development of YOLOv5-based real-time smart monitoring system for increasing lab safety awareness in educational institutions. Sensors. 22(22):8820. doi: 10.3390/s22228820.

Arya D, Maeda H, Ghosh SK, Toshniwal D, Mraz A, Kashiyama T, Sekimoto Y. 2021. Deep learning-based road damage detection and classification for multiple countries. Autom Constr. 132:103935. doi: 10.1016/j.autcon.2021.103935.

Barriera M, Pouget S, Lebental B, Van Rompu J. 2020. In situ pavement monitoring: a review. Infrastruct. 5(2):18. doi: 10.3390/infrastructures5020018.

Benjumea A, Teeti I, Cuzzolin F, Bradley A. 2021. YOLO-Z: improving small object detection in YOLOv5 for autonomous vehicles [Internet]. [accessed 2023 Nov 17]. doi: 10.48550/ARXIV.2112.11798.

Bewley A, Ge Z, Ott L, Ramos F, Upcroft B. 2016. Simple online and realtime tracking. In: 2016 IEEE International Conference on Image Processing (ICIP) [Internet]. Phoenix, AZ, USA: IEEE; [accessed 2023 Jan 9]. p. 3464–3468. doi: 10.1109/ICIP.2016.7533003.

Bochinski E, Eiselein V, Sikora T. 2017. High-speed tracking-by-detection without using image information. In: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) [Internet]. Lecce, Italy: IEEE; [accessed 2022 Dec 8]. p. 1–6. doi: 10.1109/AVSS.2017.8078516.

Boston K. 2016. The potential effects of forest roads on the environment and mitigating their impacts. Curr For Rep. 2(4):215–222. doi: 10.1007/s40725-016-0044-x.

Broström M. 2022. Real-time multi-camera multi-object tracker using YOLOv5 and StrongSORT with OSNet [Internet]. https://github.com/mikel-brostrom/Yolov5_StrongSORT_OSNet.

Brown M, Mercier S, Provencher Y. 2003. Road maintenance with opti-grade ®: maintaining road networks to achieve the best value. Transp Res Rec. 1819(1):282–286. doi: 10.3141/1819a-41.

Cao MT, Tran QV, Nguyen NM, Chang KT. 2020. Survey on performance of deep learning models for detecting road damages using multiple dashcam image resources. Adv Eng Inform. 46:101182. doi: 10.1016/j.aei.2020.101182.

Dietz P, Knigge W, Löffler H. [1984] 2011. Walderschließung: ein Lehrbuch für Studium und Praxis unter besonderer Berücksichtigung des Waldwegebaus; mit 65 Tabellen. Repr [der Ausg. Hamburg, Berlin, Parey]. Remagen-Oberwinter: Kessel.

Dodson EM. 2021. Challenges in forest road maintenance in North America. Croat J For Eng (Online). 42(1):107–116. doi: 10.5552/crojfe.2021.777.

Du Y, Song Y, Yang B, Zhao Y. 2022. StrongSORT: make DeepSORT great again [Internet]. [accessed 2022 Dec 8]. doi: 10.48550/ARXIV.2202.13514.

Erber G, Kroisleitner H, Huber C, Varch T, Stampfer K. 2021. Periodical maintenance of forest roads with a mobile stone crusher. Croat J For Eng (Online). 42(1):1–12. doi: 10.5552/crojfe.2021.862.

Ezzati S, Palma CD, Bettinger P, Eriksson LO, Awasthi A. 2021. An integrated multi-criteria decision analysis and optimization modeling approach to spatially operational road repair decisions. Can J For Res. 51(3):465–483. doi: 10.1139/cjfr-2020-0016.

Ferenčík M, Kardoš M, Allman M, Slatkovská Z. 2019. Detection of forest road damage using mobile laser profilometry. Comput Electron Agric. 166:105010. doi: 10.1016/j.compag.2019.105010.

Heidari JH, Najafi A, Alavi S. 2018. Pavement deterioration modeling for forest roads based on logistic regression and artificial neural networks. Croat J For Eng. 39(2):271–287.

Heidari MJ, Najafi A, Borges JG. 2022. Forest roads damage detection based on deep learning algorithms. Scand J Forest Res. 37(5–8):366–375. doi: 10.1080/02827581.2022.2147213.

van der Horst BB, Lindenbergh RC, Puister SWJ. 2019. Mobile laser scan data for road surface damage detection. Int Arch Photogramm Remote Sens Spatial Inf Sci. XLII-2/W13:1141–1148. doi: 10.5194/isprs-archives-XLII-2-W13-1141-2019.

Jocher G, Chaurasia A, Stoken A, Borovec J, Michael K, Fang J, Yifu Z, Wong C, Montes D, Wang Z. 2022. Ultralytics/Yolov5: V70 -

YOLOv5 SOTA Realtime Instance Segmentation [Internet]. [accessed 2022 Dec 8]. doi: 10.5281/ZENODO.3908559.

Kim YM, Kim YG, Son SY, Lim SY, Choi BY, Choi DH. 2022. Review of recent automated pothole-detection methods. Appl Sci. 12(11):5320. doi: 10.3390/app12115320.

Lanbruksdirektorat. 2016. Normaler for landbruksveier med byggebeskrivelse [standards for agricultural roads with construction description]. https://s37614.pcdn.co/wp-content/uploads/Normaler_for_landbruksveier_2016.pdf.

Luiten J, Ošep A, Dendorfer P, Torr P, Geiger A, Leal-Taixé L, Leibe B. 2021. HOTA: a higher order metric for evaluating multi-object tracking. Int J Comput Vis. 129(2):548–578. doi: 10.1007/s11263-020-01375-2.

Maeda H, Sekimoto Y, Seto T, Kashiyama T, Omata H. 2018. Road damage detection and classification using deep neural networks with smartphone images: road damage detection and classification. Comput Aided Civ Infrastruct Eng. 33(12):1127–1141. doi: 10.1111/mice.12387.

Mohiyuddin A, Basharat A, Ghani U, Peter V, Abbas S, Naeem OB, Rizwan M. 2022. Breast tumor detection and classification in mammogram images using modified YOLOv5. In: Network Asghar M, editor. Computational and mathematical methods in medicine. pp. 1–16. doi: 10.1155/2022/1359019.

MOT Challenge. [accessed 2023 Jan 3]. https://motchallenge.net/instructions/.

New Zealand forest road engineering manual. 2020. Updated February 2020. Wellington, New Zealand: NZ Forest Owners Association.

Nienaber S, Booysen MJ, Kroon RS. 2015. Detecting potholes using simple image processing techniques and real-world footage [Internet]. [accessed 2022 Dec 8]. doi: 10.13140/RG.2.1.3121.8408.

Nienaber S, Kroon RS, Booysen MJ. 2015. A comparison of low-cost monocular vision techniques for pothole distance estimation. In: 2015 IEEE Symposium Series on Computational Intelligence [Internet]. Cape Town, South Africa: IEEE; [accessed 2022 Dec 8]. p. 419–426. doi: 10.1109/SSCI.2015.69.

OpenStreetMap. OpenStreetMap [Internet]. [accessed 2023 Jan 3]. https://www.openstreetmap.org/.

Picchio R, Proto AR, Civitarese V, Di Marzio N, Latterini F. 2019. Recent contributions of some fields of the electronics in development of forest operations technologies. Electron. 8(12):1465. doi: 10.3390/electronics8121465.

Puliti S, Astrup R. 2022. Automatic detection of snow breakage at single tree level using YOLOv5 applied to UAV imagery. Int J App Earth Obser And Geoinfor. 112:102946. doi: 10.1016/j.jag.2022.102946.

Puliti S, McLean JP, Cattaneo N, Fischer C, Astrup R, Kattenborn T. 2022. Tree height-growth trajectory estimation using uni-temporal UAV laser scanning data and deep learning. An Int J For Res. 96(1):37–48. doi: 10.1093/forestry/cpac026.

Redmon J, Divvala S, Girshick R, Farhadi A. 2016. You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Las Vegas, NV, USA: IEEE; [accessed 2023 Jan 9]. p. 779–788. doi: 10.1109/CVPR.2016.91.

Roberts R, Inzerillo L, Di Mino G. 2020. Exploiting low-cost 3D imagery for the purposes of detecting and analyzing pavement distresses. Infrastruct. 5(1):6. doi: 10.3390/infrastructures5010006.

Saarenketo T, Aho S. 2005. Managing spring thaw weakening on low volume roads [Internet]. Inverness, Scotland: The Highland Council, Transport, Environmental & Community Service. https://www.roadex.org/wp-content/uploads/2014/01/2_3-Spring_Thaw_Weakening_l.pdf

Sattar S, Li S, Chapman M. 2018. Road surface monitoring using smartphone sensors: a review. Sensors. 18(11):3845. doi: 10.3390/s18113845.

Starke M, Geiger C. 2022. Machine vision based waterlogged area detection for gravel road condition monitoring. Int J For Eng. 33(3):243–249. doi: 10.1080/14942119.2022.2064654.

Susnjar M, Pandur Z, Nevecerel H, Lepoglavec K, Bacic M. 2020. Development of a new method for assessing condition of forest road surface. JCE. 71(12):1121–1128. doi: 10.14256/JCE.2462.2018.

Talbot B, Pierzchala M, Astrup R. 2017. Applications of remote and proximal sensing for improved precision in forest operations. Croat J For Eng. 38(2):327–336.

Waga K, Tompalski P, Coops NC, White JC, Wulder MA, Malinen J, Tokola T. 2020. Forest road status assessment using airborne laser scanning. Forest Sci. 66(4):501–508. doi: 10.1093/forsci/fxz053.

Zhou K, Yang Y, Cavallaro A, Xiang T. 2019. Omni-scale feature learning for person re-identification [Internet]. [accessed 2022 Dec 8]. doi: 10.48550/ARXIV.1905.00953.