Journal of Outdoor Recreation and Tourism 35 (2021) 100387

Contents lists available at ScienceDirect

Journal of Outdoor Recreation and Tourism

journal homepage: www.elsevier.com/locate/jort

Research Article

Comparing established visitor monitoring approaches with triggered trail camera images and machine learning based computer vision

Jeroen Staab [a,b,*], Erica Udas [c,d], Marius Mayer [e,f], Hannes Taubenböck [a,g], Hubert Job [g]

[a] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Germany
[b] Humboldt-University Berlin, Geography Department, Germany
[c] International Center for Integrated Mountain Development, Katmandu, Nepal
[d] University Greifswald, Institute of Botany and Landscape Ecology, Germany
[e] University Greifswald, Institute of Geography and Geology, Germany
[f] University Innsbruck, Department of Strategic Management, Marketing andTourism, SME and Tourism, Austria
[g] University of Würzburg, Institute of Geography and Geology, Germany

ARTICLE INFO

Keywords:
Visitor monitoring
Computer vision
Convolutional neural network
Camera
Protected areas

ABSTRACT

The management of protected areas and other recreational landscapes is subject to a variety of challenges. One aspect hereof, visitor monitoring, is crucial for many management and valuation tasks of ecosystem services. Its core data are visitor numbers which are costly to estimate in absence of entry fees for protected areas or recreational landscapes. Camera-based approaches have the potential to be both, accurate and deliver comprehensive data about visitor numbers, types and activities. So far, camera-based visitor monitoring is, however, costly due to time consuming manual image evaluation. To overcome this limitation, we deployed a convolutional neural network and compared its hourly counts against existing visitor counting methods such as manual in-situ counting, a pressure sensor, and manual camera image evaluations. Our study is the first one to implement, and explicitly assess the performance of a computer vision approach for visitor-monitoring. The results showed that the convolutional neural network derived comparable visitor numbers to the other visitor counting approaches regarding visitation patterns and numbers of visits. Further, our approach also allowed for counting dogs and recreational equipment such as backpacks and bicycles in automatic manner. We thus conclude that it is a fast and reliable method that could be used in protected areas as well as in a much wider array of visitor counting settings in other recreational landscapes.

Management implications: Managers of protected and recreational areas could benefit from our comparisons of convolutional neural network camera image evaluations with existing visitor counting approaches as:

- Time-consuming manual image evaluation can be replaced by computer vision approaches based on convolutional neural networks (40 h to manually analyze more than 13,000 images by one expert vs. 10 h to do it automatically in the background).
- In contrast to pressure sensors, this approach also allows to differentiate visitor types and activities (dog-walking, cycling, etc.) at comparably low-costs.
- Future efforts should concentrate on training specific convolutional neural networks dedicated to visitor monitoring in recreational settings which could process imagery at real-time in the field using single-board computers.
- Nevertheless, this approach is prone to the usual disadvantages of camera-based visitor monitoring (risks of theft, vandalism, malfunctioning; data security issues), which need to be considered when setting up the device.

* Corresponding author. German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Münchener Str 20, 82234 Wessling, Germany.
E-mail addresses: Jeroen.Staab@dlr.de (J. Staab), Erica.Udas@icimod.org (E. Udas), Marius.Mayer@uibk.ac.at (M. Mayer), Hannes.Taubenboeck@dlr.de (H. Taubenböck), Hubert.Job@uni-wuerzburg.de (H. Job).

https://doi.org/10.1016/j.jort.2021.100387
Received 1 August 2020; Received in revised form 4 November 2020; Accepted 21 April 2021
Available online 10 May 2021

1. Introduction

Protected areas are a major destination for outdoor recreation (Balmford et al., 2015; Schägner et al., 2018). Along with fostering physical health and mental well-being (Rathmann et al., 2020; Taye et al., 2019), outdoor recreation also generates considerable economic impacts and job opportunities (Highfill & Franks, 2019; Job, Merlin, Metzler, Schamel, & Woltering, 2016; Mayer et al., 2010). Although the importance of outdoor recreation is generally well recognized, when it comes to concrete land-use decision making, it is necessary to provide clear evidence of an area's relevance to recreational services (and other related ecosystem services). In this context, accurate visitor numbers have been concluded to be the most important parameter in economic impact and recreational value modelling of protected areas (Woltering, 2012; Mayer & Woltering, 2018; Schägner et al., 2017, 2018). Further, a comprehensive visitor monitoring system is relevant for developing personnel deployment plans, targeted visitor information and marketing measures (Arnberger et al., 2016; Ziesler & Pettebone, 2018), to improve recreation opportunities (Ankre et al., 2016), to assess the ecological carrying capacities of protected areas and recreational landscapes (Cessford & Muhar, 2003; Ziesler & Pettebone, 2018), to analyze the influence of weather, exchange rate and media coverage variations (Millhäusler et al., 2016) as well as for informed visitor management avoiding potential conflicts (Ankre et al., 2016; Job, Schamel, & Butzmann, 2016; Rupf et al., 2014) and crowding (Schamel & Job, 2013). However, in most European countries like Germany or Sweden and elsewhere e.g. New Zealand too, there are open access policies governing protected areas, so that, for example, visitor data cannot be derived from entrance fees (Hannemann & Job, 2003). Therefore, a plethora of visitor monitoring approaches like field observations and mechanical counters exist (see review by Cessford & Burns, 2008).

Nevertheless, all approaches have specific pros and cons, which must be considered according to the individual case, local circumstances (e.g. pathway layout), legal requirements, financial restrictions and methodological competences (Hornback & Eagles, 1999; Muhar et al., 2002; Cessford & Muhar, 2003; Kajala et al., 2007; Arnberger, 2007; Spenceley et al., 2021 for a recent overview). For example, direct observation by trained staff in the field can distinguish walking directions, recognize sport equipment and can be combined with conducting interviews. However, as human resources are expensive, full-time counting is practically impossible and thus, visitor monitoring observations are usually sampled and extrapolated (Job et al., 2005; Mayer et al., 2010). In comparison, automated approaches such as the use of photo-electronic or pressure sensitive sensors, which are widely used for long-term visitor monitoring, have economic advantages (Arnberger, 2007; Arnberger et al., 2016; Cessford & Muhar, 2003). However, these approaches cannot differentiate user groups based on recreational equipment. Furthermore, the sensors cannot discriminate between humans and wildlife, leading to suspiciously high numbers e.g. during the hunting season (Hodges, 2009).

In this trade-off between operational costs and visitor monitoring comprehensibility, Arnberger et al. (2005) compared time-lapse video observations evaluated by a human observer against direct field observations. With technical improvements and decreasing prices, commercially available (surveillance) cameras can be utilized in the domain of managing protected areas and other recreational landscapes (Arnberger, 2007; Cessford & Muhar, 2003). These devices differ in image resolution, sensor type (usually optical) and operational mode, e.g. continuous video recordings, time-lapse photography (Arnberger et al., 2005; Kahler & Arnberger, 2008) or motion-triggered recordings (Czachs & Brandenburg, 2014; Lupp et al., 2016; Miller et al., 2017). The deployment of the latter is most common, as triggered trail cameras are familiar to staff working in protected areas due to their use in wildlife monitoring (Kays et al., 2011).

However, using cameras in this domain still has some practical limitations. First, theft and vandalism may occur (Czachs & Brandenburg, 2014). Secondly, ethical aspects and the respective legal frameworks need to be considered. In Germany for example, among other regulations, a signage is required to inform visitors about ongoing camera observations, data needs to be destroyed at the end of the project, and individuals may not be identified at any time (BDSG, 2018). Third, and most important regarding long-term monitoring projects, the manual evaluation still requires large amounts of human resources and time. Therefore, this approach is not feasible in many cases.

In 1995 MMuhar, Zemann, & Lengauer tried to tackle the latter limitation, by utilizing early computer vision techniques. Nevertheless, manual camera data interpretation is still the most common evaluation approach, although meanwhile, computer scientists have developed precise and fully automated image interpretation methods endowing automated cars with computer vision to avoid hitting pedestrians (Dollár et al., 2012; Zhang et al., 2017). Among such technologies, deep convolutional neural networks (CNN) are the most capable (Brunetti et al., 2018) and are also often deployed in environmental sciences too (e.g. Stiller et al., 2019; Wurm et al., 2019). Mechanically, such deep learning methods consist of multiple automatically feature engineering layers, which, provided that the training data is adequate, have proven to make very accurate predictions (Guo et al., 2016). Particularly in the domain of wildlife monitoring, and impelled by the *Snapshot Serengeti* dataset (Swanson et al., 2015), CNNs are often combined with triggered trail cameras (Falzon et al., 2020; Gomez Villa et al., 2017; Schneider, Taylor, & Kremer, 2018; Yousif et al., 2019) and have filled the gap between capturing image data in the field and analyzing it for management decision making. Notably, Yousif et al. (2019) deployed the method to particularly exclude pedestrians along with blank images triggered by moving vegetation in their ecological investigations. However, to the best of our knowledge, no such technologies have been deployed to explicitly count recreational visitors so far – the extensive reviews by Pickering et al. (2018) and Ziesler and Pettebone (2018) do not mention them, for instance.

This paper introduces such a CNN to explicitly improve automatic visitor counting in outdoor recreation settings. We elaborate on how it competes against conventional approaches like the manual analysis of camera pictures and other established visitor counting methods. Our extensive assessment thus provides both researchers and practitioners with a possibility of making informed choices about which visitor monitoring approach to implement and for what kind of outdoor recreational settings, considering costs, efforts, and expected results.

This paper is structured as follows: After a presentation of our research area (section 2) and instruments utilized for visitor monitoring and different image evaluation methods (section 3), we first present the results of each counting approach individually, then compare the observations against each other, and finally present the annual cumulated number of visits counted by each approach at last (section 4). In section 5, we summarize our results and discuss the advantages and limitations of the visitor monitoring approaches, and conclude their implications for future visitor monitoring in and outside protected areas in section 6.

2. Study site

Eldena Forest Nature Reserve (EFNR) is located at the southeastern rim of the city of Greifs-wald , Germany. The forest is owned by Greifswald University and is spread over 411 ha, including open spaces, forest trails and paths. It was designated as a nature protection area in 1961. Some small forest patches (29 ha) consist of old growth oak and beech forests which are under strict protection without any management activities. Given its close vicinity to the city of Greifswald, EFNR is frequently visited by local residents, students, and tourists for recreation, study and other purposes; however, exact visitor numbers have never been estimated so far. This information would be very relevant for forest management planning as recreation is one of the major forest ecosystem services. Its valuation would also allow for the assessment of potential tradeoffs and synergies with other ecosystem services (Udas et al., 2018). There are seven major entry points into EFNR, and based on the communication with local foresters, the entries at A, B, C and D are the ones mostly used due to their proximity to residential areas of Greifswald (Fig. 1).

While conducting systematic visitor monitoring in EFNR, three different visitor counting methods were deployed at different entrances in 2015. At all seven entrances, manual in-situ visitor counting was carried out. In addition, at the most frequented entrances like at entrance A, a pressure sensor was installed, and triggered trail cameras were installed at entrances B, C, D and E.

3. Methodology

3.1. Manual counting

As a fundamental benchmark, manual in-situ visitor counting was conducted following a visitor counting method used in many protected areas (Job et al., 2005, Job, Merlin, et al., 2016; Job, Schamel et al., 2016; Job, 2008; Mayer et al., 2010; Woltering, 2012, Rein & Baláš, 2015; Rein et al., 2019). The manual in-situ visitor counting was scheduled for a total of twelve days distributed throughout the year 2015. We considered different weather conditions, public holidays, weekends and weekdays, and a combination of which gave four day-type categories. On each census day, trained students counted visitors at each entrance for 10 h between 09:00 and 19:00. To avoid double counting from different entrances, only visitors entering the forest were counted. There were, however, a few missing field observations at entrances A, E, F and G because of logistical reasons. These gaps were substituted based on statistical relationships (see Ziesler & Pettebone, 2018), i.e. the percentage share and relation of visitor counts at other entrances for the same day or the same day-type category.

To determine the annual cumulative number of visits in EFNR, the actual data from the sampled days were extrapolated. The extrapolation procedure accounted for seasonality, weekends/weekdays and the weather situation and is a standard procedure used by many studies (Job et al., 2005, Job, Merlin, et al., 2016, JJob, Schamel et al., 2016; Job, 2008; Mayer et al., 2009, 2010; Woltering, 2012, Rein & Baláš, 2015; Rein et al., 2019). Daily weather data of Greifswald in 2015 was obtained from the online database portal of the German Meteorological Service. Based on this data, Z-standardized values were estimated for daily mean air temperature (T_z), daily sunshine hours (S_z), and daily precipitation (P_z). Referring to a moving window of 15 days before and after the respective days, positive values of $\frac{1}{3}(T_z + S_z - P_z)$ were considered to be good weather, whereas negative values were considered as bad weather (Mayer et al., 2009, 2010). For official public holidays in 2015, the holiday calendar of the Federal State of Mecklenburg-Western Pomerania was considered. Table 1 depicts the different combinations of good and bad weather along with holidays and working days condensed into four day-type categories and their respective frequency in 2015. Finally, extrapolation of the number of visits for the missing dates was done by using a categorical linear regression without intercept. The $\beta_{day-type}$ estimates correspond to the mean value of visitor count per day-type category.

Eventually, as the manual in-situ visitor data for the sampled days

Table 1
Categorization of the four day-types at EFNR in 2015.

Day-type	Combination			N
I	Good weather	+	weekend/public holiday	56
II	Good weather	+	working day	136
III	Bad weather	+	weekend/public holiday	55
IV	Bad weather	+	working day	118
Sum				365

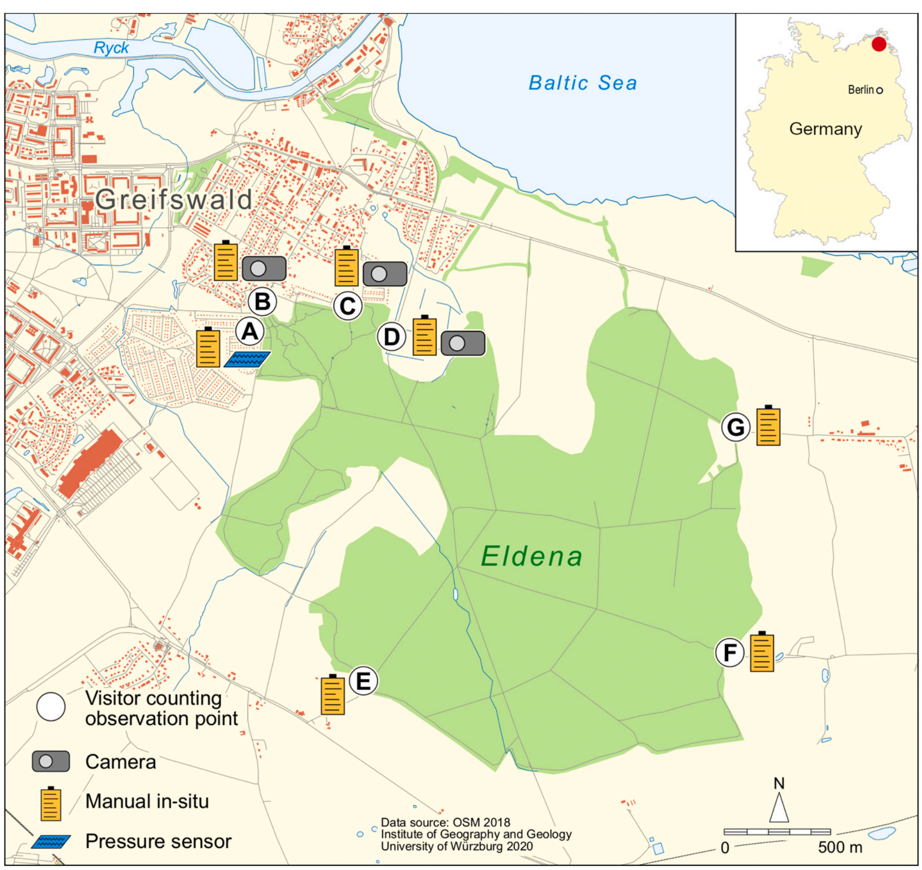

Fig. 1. Map of Eldena forest nature reserve.

was collected from 09:00 until 19:00, a supplement visitor number for the remaining hours was added as a percentage fraction of visitors entering the forest outside the field observation window estimated from the automated counters installed at entrances A, B, C, and D (increments presented in Tables 3 and 4).

3.2. Pressure sensor

Given the limited human resources for manual counting, the seasonal variation of visitation was captured with an automatic counting machine. An Ecocounter® pressure sensor was installed throughout the year 2015 at entrance A, one of the most frequented entrances suitable for the pressure sensor due to its narrow path. Whenever a person steps on it - according to its manufacturer it can detect weight fluctuations in a range of 10 kg to 3,5t - this is registered. Since the pressure sensor is able to detect bi-directional counts, both in- and outgoing visitor data was recorded on an hourly basis for each day (24 h). Although it is common to calibrate automated visitor counting (Pettebone et al., 2010), we explicitly refrained from calibrating the results allowing for rigid comparisons of the raw observations. The installation of the pressure sensor took four persons about 1.5 h. After that, the sensor worked incessantly and uploaded its data automatically to the digital cloud.

3.3. Triggered trail cameras

3.3.1. Hardware set-up

Triggered trail cameras were installed in January 2015 at entrance B and in March 2015 at entrances C, D and E to automatically capture the visitors entering EFNR. Camera B was installed within the EFNR at a place diagonal to a forest path. Camera C was installed overlooking an intersection at the very entrance to EFNR, whereas camera D was installed along a narrow single trail leading to the EFNR. In all cases we chose the lowest image resolution and the motion-triggered defaults used for wildlife monitoring. Whenever a warm object moved within the infrared-detection field, an image was taken. In order to avoid vandalism (Arnberger et al., 2005; Czachs & Brandenburg, 2014), the cameras were hidden in bird boxes at a height of about 4 m in trees. Nevertheless, the camera at entrance E was stolen shortly after its installation whereas the cameras at entrances C and D were turned away from their original positions regularly, and eventually both devices were stolen. As a consequence, there was occasionally no imagery available for these entrances. Subsequently, the camera at entrance B had technical problems in mid-June 2015 and was replaced at the end of September 2015, resulting in about three months of data discontinuity. The overall sufficing observations for the missing days were eventually filled using the same annual extrapolation method defined earlier.

Since this study compared manual image evaluation and automated image analyses regarding costs and performance, the archived images were first scrutinized to allow a direct comparison of the methods and reduce the amount of data. Using the Windows native Photo Viewer, images, that did not show any visitors (i.e. that captured animals) or those that were redundant (i.e. two consequent images showing the same person) were excluded from the analysis.

3.3.2. Manual image evaluation

The images for monitoring visitors in EFNR were first evaluated manually. Fig. 2 illustrates three semantics of image interpretation to estimate visitor numbers. *Semantic 1* only includes actual persons visible on an image. In images, where people conceal each other while walking in staggered position or chipped at the edge of an image, they were counted as long as at least any body part was identifiable. In *Semantic 2* the same rules are applied, but in addition, also contextual decisions were made. Assuming for example that at least one child is inside a baby-stroller, this was considered a visitor. However, if a small child was seen walking beside the baby-stroller, then only the visible child was counted. *Semantic* 3 is similar to *2*, but walking direction are distinguished as well. In this case, only visitors entering the EFNR were counted. To understand the impact of these semantics on the visitor count, a sample of 250 reference images were manually evaluated per entrance at B, C and D.

The counting scheme of *Semantic 3* contains most information relevant to the domain of visitor monitoring. Therefore, following a conventional approach (Czachs & Brandenburg, 2014; Kahler & Arnberger, 2008; Miller et al., 2017), this semantic was applied to the entire set of archived images. To minimize personal observational biases, the visual analysis was conducted by a single person. When there were big groups visiting EFNR, a pair of one or two sequential images was carefully evaluated to make sure that all visitors are counted. The number of visitors entering the EFNR was documented on an hourly basis for each day in a spreadsheet.

3.3.3. Automated image analyses

As investigated by Zhang et al. (2017), advanced computer vision technologies such as deep CNNs can detect pedestrians at very high accuracies. To do so, many different network architectures for analyzing still imagery and corresponding training utilities are available (Brunetti et al., 2018). The design and training of a new network from scratch, however, requires appropriate training data, specific hardware, and expertise. For the purposes of this research, we therefore considered a pre-trained image analyzing framework. The most important decision drivers were computational performance and a broad training capable of differentiating multiple user groups. One such tool was developed by Redmon, Divvala, Girshick, & Farhadi (2016), respectively Redmon & Farhadi (2017), 2018) and, as its name indicates, *You Only Look Once* (YOLO) is very fast in grasping an image's content. Therefore, it can be deployed in real-time applications (e.g. Han et al., 2020). As a result of the versatile training data, the pre-trained algorithm detects several object classes in an image, among which are persons, bicycles, backpacks and dogs – categories of special interest to characterize visitors in recreational landscapes.

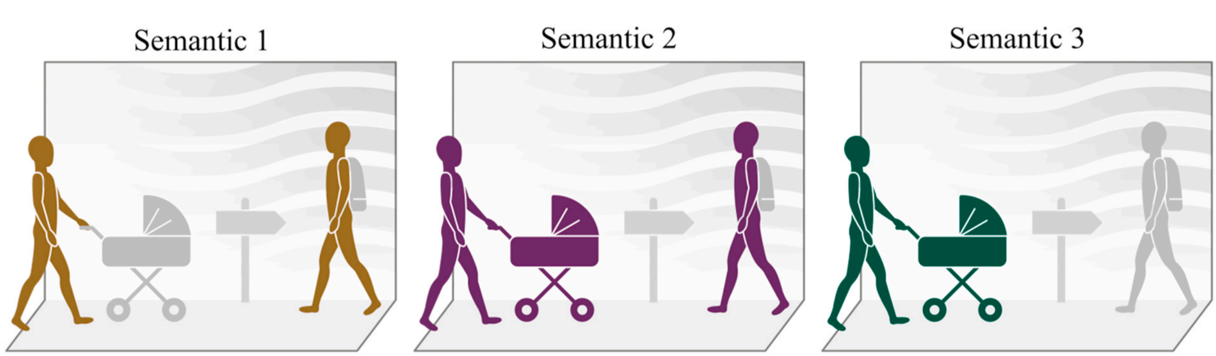

Fig. 2. Illustration for manual image counting Semantics 1–3. Counted persons are colored. The person on the right side, carrying a backpack, is considered as leaving the EFNR.

When deploying YOLOv3 (Redmon & Farhadi, 2018) on an image, each detected object is returned with its consequent description. This includes the object's location (i.e. a bounding box), a *class label* describing the object semantically (e.g. "person" or "car") and finally the *certainty* as an indicator of the network's classification confidence (e.g. 0.92 = 92%). Eventually, the detections are filtered according to their *certainty* exceeding a predefined threshold. As this threshold could be compared with the minimum weight of a pressure sensor, naturally, this sensitivity threshold is inversely correlated with the number of detected visitors. In order to select an appropriate threshold, which by default is $\geq$ 0.25, a sensitivity analysis was conducted with 0.05 increments. Therefore, the semantically labeled images were used as reference data. However, to rule out semantic biases, we restricted this assessment to images where *Semantic 1*, *2*, and *3* agreed and referred to this subset as *Semantic 0*. Instances where the automatic approach underestimates the reference data are defined as false negatives. Vice versa, a false positive (an object wrongly labeled as "person") is defined as overestimation. Lastly, the errors were assessed together with the fraction of correctly evaluated images. We expect that this extensive assessment helps to better understand which semantic is coded in the pre-trained CNN (Redmon & Farhadi, 2018).

Lastly, the processing time of YOLO in an operational setup was measured on different platforms. To outline the plethora of potential applications, three distinct computers were used. Representing a small and local use-case, we used a low-cost microprocessor (*RaspberryPi 3+*). While the resources of this single board computer are very limited, we assume it could be used for ad-hoc image evaluations in the field. In the second case, we used a common laptop (*Lenovo T460*). Such and similar devices are used by many research groups and can be utilized for small or experimental projects. Last but not least, as we processed the complete image archive of the study together, we benchmarked the processing time on a high-performance computer with a dedicated graphical processing unit (*NVIDIA P100*). Here, we acknowledge that such equipment may only be available to scholars in the field of computer vision and artificial intelligence. However, for large-scale projects, with cameras streaming their data to such a central server, this may allow for processing multiple images in real-time.

3.4. *Comparison of counting approaches*

Eventually we directly compared the results of all counting approaches. For each entrance the raw, hourly results per counting approach were set against each other using R and Excel. The statistical deviations were measured using Pearson's correlation and a linear model without intercept.

4. Results

4.1. *Outcome of the individual counting approaches*

4.1.1. *Manual in-situ observations*

In total, more than 500 h of manual observations were conducted in the field, and a total of 1768 visits were counted entering the EFNR. As shown in Table 2, overall, the linear extrapolation based on day-types explains large parts of the observed variance (Ø adj. $R^2 = 0.780$). The visitation estimates for day-type I and III are significant for all entrances. For similar weather conditions, the average number of visits is lower during working days (except for G, bad weather). However, for good weather conditions the coefficients' standard error is larger on working days. That said, here the number of visits may even vary by the same magnitude as the estimated overall count. Regarding the estimates for the same type of working day, and respective weekends/holidays, it appears unexpected that on average slightly fewer visits occur at EFNR during good weather conditions. However, this is not the case on weekend/holidays at entrances F and G. Also, this trend cannot be confirmed on working days at entrance C, where approximately 44.4% more visits occur during good weather.

4.1.2. *Pressure sensor*

Over the course of 2015 (see Fig. 3), in total 21,912 visits were detected entering the EFNR as opposed to 20,439 visits leaving it (51.7% vs. 48.3%). The seasonal variations in visitation of EFNR are pronounced

Table 2
Visits counted manually from 09:00–19:00. a) Yearly extrapolation model, depicting average number of visits aggregated per day type category. b) Total extrapolated number.

	a)					b)
	Average good weather		Average bad weather		Adj. R^2 [MAE]	Total Visits (between 09:00–19:00)
	weekend/PH	working day	weekend/PH	working day		
	β_I	β_{II}	β_{III}	β_{IV}		
A	81.412 (18.65) **	23.104 (22.84)	116.841 (18.65) ***	61.587 (16.15) **	0.853 [20.385]	21,395
B	54.839 (14.24) **	20.042 (17.44) 61.667	(14.24) **	32.000 (12.33) *	0.758 [17.409]	12,964
C	33.333 (7.378) **	32.500 (9.049) **	49.000 (7.389) ***	22.500 (6.399) **	0.877 [8.472]	11,637
D	23.851 (7.759)	14.000 (9.503)	27.013 (7.759) **	15.697 (6.719) *	0.677 [7.771]	6,578
E	20.968 (10.774) #	6.236 (13.195)	37.982 (10.774) **	13.841 (9.331)	0.550 [9.748]	5,745
F	15.641 (2.073) ***	4.735 (2.539) #	12.500 (2.073) ***	9.391 (2.073) ***	0.909 [2.163]	3,315
G	24.108 (4.836) **	10.907 (5.922)	17.982 4.836) **	20.317 (4.188) **	0.837 [5.874]	6,220

The numbers within the brackets indicate (): standard error and []: mean average error; PH: Public holiday; Levels of significance: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, #$p < 0.1$.

Table 3
Visits detected by pressure sensor. a) Yearly extrapolation model, depicting average number of visits aggregated per day type category. b) Total absolute number. c) Share of visits outside the manual in-situ observation window.

	a)					b)	c)
	Average good weather		Average bad weather		Adj. R^2 [MAE]	Total	Share of visits 19:00–09:00
	weekend/PH	working day	weekend/PH	working day			
	β_I	β_{II}	β_{III}	β_{IV}			
A_{In}	104.411 (5.453) ***	59.382 (3.499) ***	63.582 (5.502) ***	38.144 (3.756) ***	0.709 [30.195]	21,921	12.2%
A_{Out}	88.338 (4.688) ***	55.360 (3.008) ***	60.691 (4.730) ***	39.195 (3.229) ***	0.733 [25.605]	20,439	15.4%

The numbers within the brackets indicate (): standard error and []: mean average error; PH: Public holiday; Levels of significance: ***p < 0.001, **p < 0.01, *p < 0.05, #p < 0.1.

Table 4
Visits counted with Camera^{Manual}. a) Yearly extrapolation model, depicting average number of visits aggregated per day-type category. b) Total extrapolated number. c) Share of visits outside the manual in-situ observation window.

	a)						b)	c)
		Average good weather		Average bad weather		Adj. R^2 [MAE]	Total	Share of visits 19:00–09:00
		weekend/PH	working day	weekend/PH	working day			
	N	β_I	β_{II}	β_{III}	β_{IV}			
B	208	47.629 (5.879) ***	34.438 (4.071) ***	45.548 (6.247) ***	25.493 (4.187) ***	0.518 [23.894]	12,864	14.8%
C	81	52.400 (4.276) ***	21.387 (2.429) ***	35.813 (3.381) ***	14.292 (2.760) ***	0.817 [9.571]	9,499	11.5%
D	55	39.889 (3.178) ***	14.381 (2.080) ***	23.111 (3.178) ***	7.562 (2.383) **	0.828 [6.195]	6,353	14.7%

The numbers within the brackets indicate (): standard error and []: mean average error; PH: Public holiday; Levels of significance: ***p < 0.001, **p < 0.01, *p < 0.05, #p < 0.1.

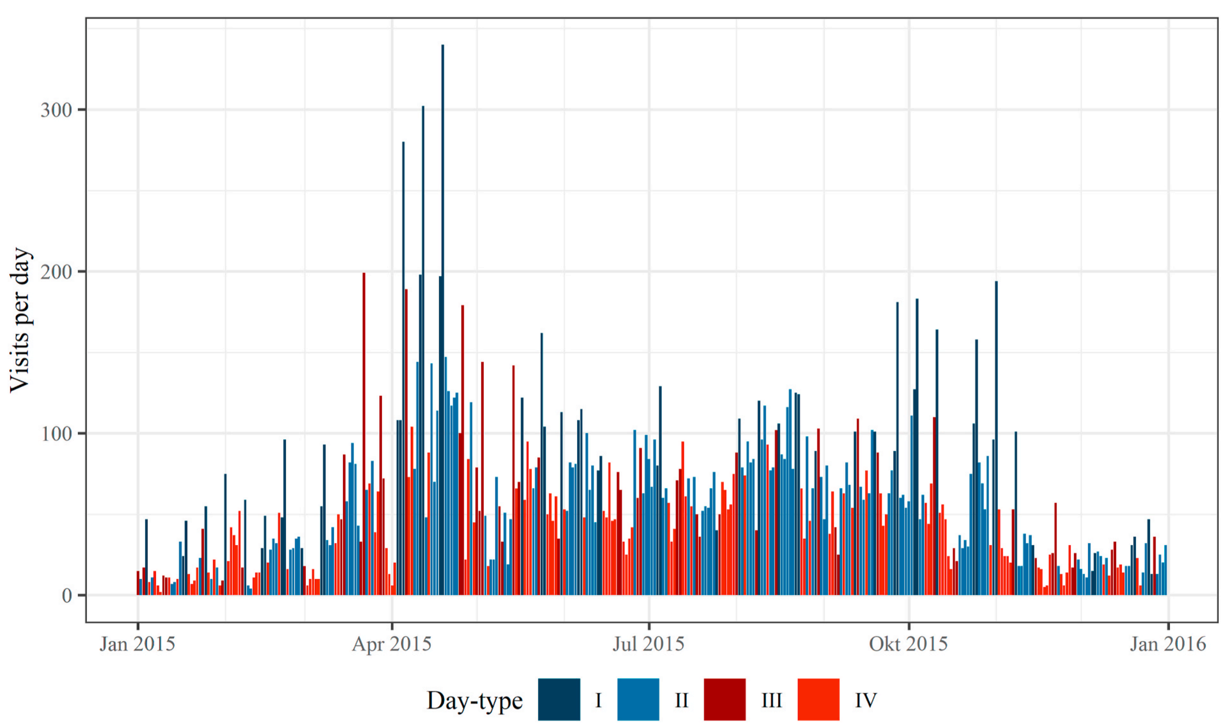

Fig. 3. Visits per day detected by pressure sensor at entrance A.

with two culmination periods in April and October concurring with blossoming and blooming in spring and leaf coloration in autumn, respectively. Midsummer, however, shows lower visitation, while winter months have the lowest overall frequentation. Further, regarding the missing observation time window during manual counting, we measured a fraction of 12.2% entries between 19:00 and 09:00 at entrance A.

Although the pressure sensor worked flawlessly for 365 days, for a better comparison with the other counting approaches, we present the average number of visits per day type by fitting a linear extrapolation model. At this point, it is interesting to stress that Table 3 shows similar visitation patterns with regard to equal weather conditions as the ones described above. However, for similar types of working day, and respective weekend/holiday, the frequency of visitation registered by the pressure sensor is higher for good weather. Besides, Table 3 also shows that the rigid extrapolation model is capable of explaining 70.9% of the overall variance with a mean average error (MAE) of approximately 30 visits.

4.1.3. Camera data

4.1.3.1. Manual evaluation. Three different counting schemes were applied to assess the impact of various counting semantics. Using a sample of 250 images from entrances B, C and D each, a total of 1,150, 1179 and 1153 persons were counted using *Semantics 1*, *2*, or *3* respectively. Overall, the counts agreed in 90.1% of the images and the corresponding correlation coefficients between the three image interpretation semantics are very high. Nevertheless, these coefficients stress the semantic details as well. Pearson r is highest between *Semantic 1* and *2* ($r_{12} = 0.984$) and tends to be lower for *Semantic 3*, which exclusively focused on persons entering the EFNR ($r_{13} = 0.947$; $r_{23} = 0.944$). The latter two values reflect that *Semantic 2* and *3* share their definition of visitors (i.e. expect a child inside a baby-stroller), while *Semantic 1* literally counted persons only.

Regarding the interests of visitor monitoring in outdoor recreation such as avoiding double counting, *Semantic 3* was applied on the complete image archive. In total, a single researcher devoted circa 40 h counting visitors entering EFNR. Thereby, in total 7,352, 2103 and 900 visits were manually counted at entrances B, C, and D respectively. As the number of valid observation days was limited (see N in Table 4) due to technical issues, vandalism and theft, deriving linear extrapolation models based on day-types I-IV was obligatory again. When investigating the visitation patterns, although less stark, the same plausible trends as observed by the pressure sensor are confirmed.

4.1.3.2. Automated approach. Our goal is to assess the potential to automate camera evaluations at comparable precision. First, we investigated YOLO's detection certainties and identified an appropriate threshold using the semantically labeled reference images. However, to rule out semantic biases, we also conducted sensitivity analyses with the 90.1% samples belonging to *Semantic 0* (N = 682). Looking at Fig. 4a, it is clear that this dataset systematically outperforms *Semantic 1–3*. After starting at a *certainty* threshold ≥0.05 (where on average 9.410 false positives are detected per image, see Fig. 4b), we initially see a steep increase in accuracy. At thresholds ranging between ≥0.15 and ≥0.25, we see that up to 92% of the reference images were evaluated correctly, before a slow decrease is measured again. Eventually, when requiring YOLO's predictions to be absolutely confident (*certainty* = 1), overall 60% of the reference images were interpreted correctly (as here on average 0.526 false negatives occur, see Fig. 4b). At a threshold of ≥0.1, however, YOLO manages to balance best between detecting neither too few, nor too many people (0.050 false negative and 0.041 false positives per image respectively, see Fig. 4b). Hence, based on these considerations we set the sensitivity threshold to ≥0.1 prediction *certainty*, for evaluating of the complete data set deploying the CNN.

Considering the 13,377 images at entrances B, C and D, a total number of 18,536 persons were detected. Then we applied the annual extrapolation method at each entrance (see Table 5a). Again, all p-values of the linear coefficients are highly significant (<0.01). However, when comparing the mean visit count per day-type directly against the other models presented above, it is crucial to bear in mind that YOLO did not differentiate walking directions. As a consequence, these numbers

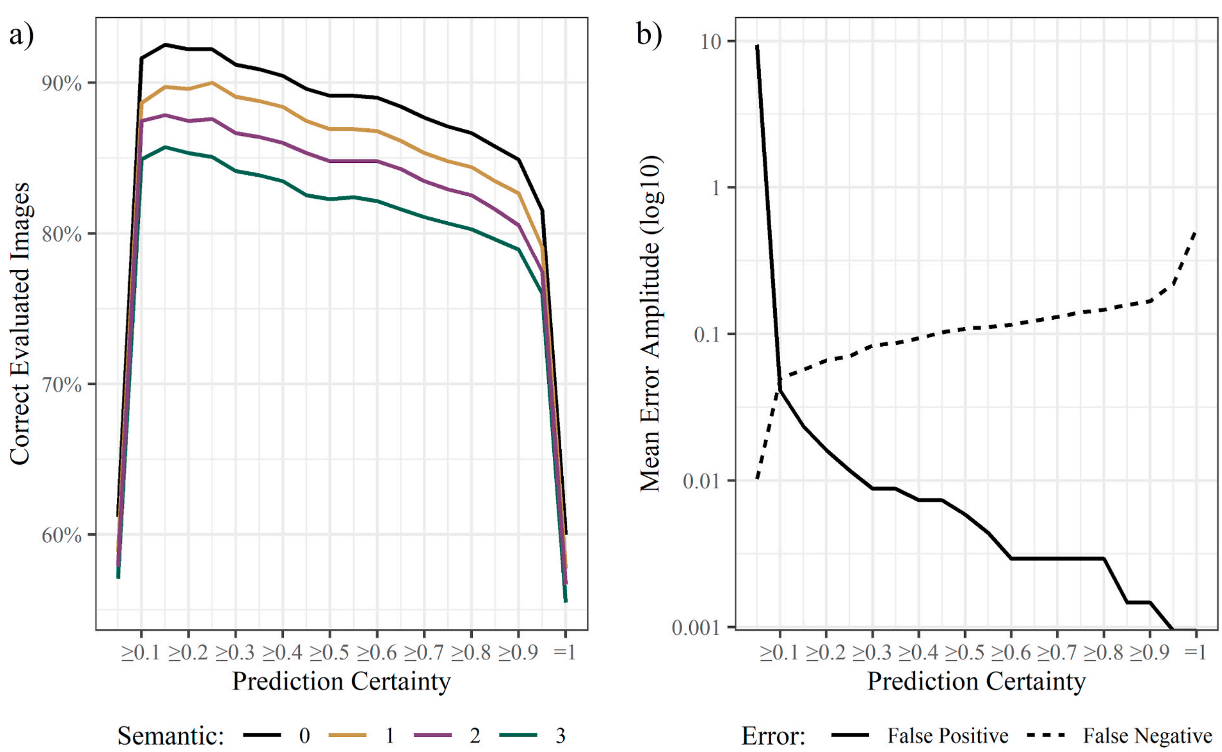

Fig. 4. Sensitivity analyses using YOLO and manually labeled images. a) Percentage of correct evaluated images regarding semantic 0–3. b) Mean amplitude of error.

Table 5
Visits counted with Camera^{YOLO}. a) Yearly extrapolation model, depicting average number of visits aggregated per day-type category. b) Total extrapolated number.

	a)						b)
		Average good weather		Average bad weather		Adj. R^2 [MAE]	Total
		weekend/PH	working day	weekend/PH	working day		
	N	β_I	β_{II}	β_{III}	β_{IV}		
B	229	62.429 (9.630) ***	50.878 (6.892) ***	66.569 (11.032) ***	40.068 (7.304) ***	0.410 [42.767]	18,799
C	81	114.600 (8.438) ***	52.323 (4.793) ***	73.375 (6.671) ***	31.042 (5.447) ***	0.848 [18.580]	21,232
D	55	87.222 (6.130) ***	25.909 (3.921) ***	44.111 (6.130) ***	13.813 (4.598) **	0.844 [11.766]	12,464

The numbers within the brackets indicate (): standard error and []: mean average error; PH: Public holiday; Levels of significance: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, #$p < 0.1$.

include entering as well as leaving visitors. The relative visitation patterns however align (except lower visitation during good weather on weekends/public holidays at entrance B).

For a better understanding of how the CNN interprets the images, Fig. 5 illustrates some common scenarios. Although at the cost of positional errors and detection *certainty*, the two more challenging examples (Fig. 5c and d) show that YOLO is capable of distinguishing visitors walking in staggered position. Further, the pre-trained network also counted objects important in the context of visitor monitoring in protected areas: 2380 bicycles, 2336 backpacks and, among other categories, 1839 dogs. Interestingly, the share of particular objects per image varies between the entrances (see Table 6). For example, parked cars were highest at entrance C and the share of bicycles lowest at entrance D, where the narrow trail with protruding roots reduces the comfort of cycling.

Table 6
Objects per image counted with Camera^{YOLO} at each entrance.

	Person	Backpack	Handbag	Bicycle	Dog	Car
B	1.45	0.20	0.13	0.45	0.15	0.01
C	1.40	0.17	0.14	0.51	0.14	0.24
D	1.40	0.10	0.07	0.21	0.12	0.00

Lastly, we present the number of images evaluated per minute using YOLO on three different processing platforms, as this information might be relevant for future studies. The single board computer (*RaspberryPi 3+*) analyses 1.2 images per minute. Using a laptop (*Lenovo T460*) is approximately three times faster (3.3 images per minute). The high-performance computer with a dedicated graphical processing unit (*NVIDIA P100*) was capable of processing 21.9 images per minute.

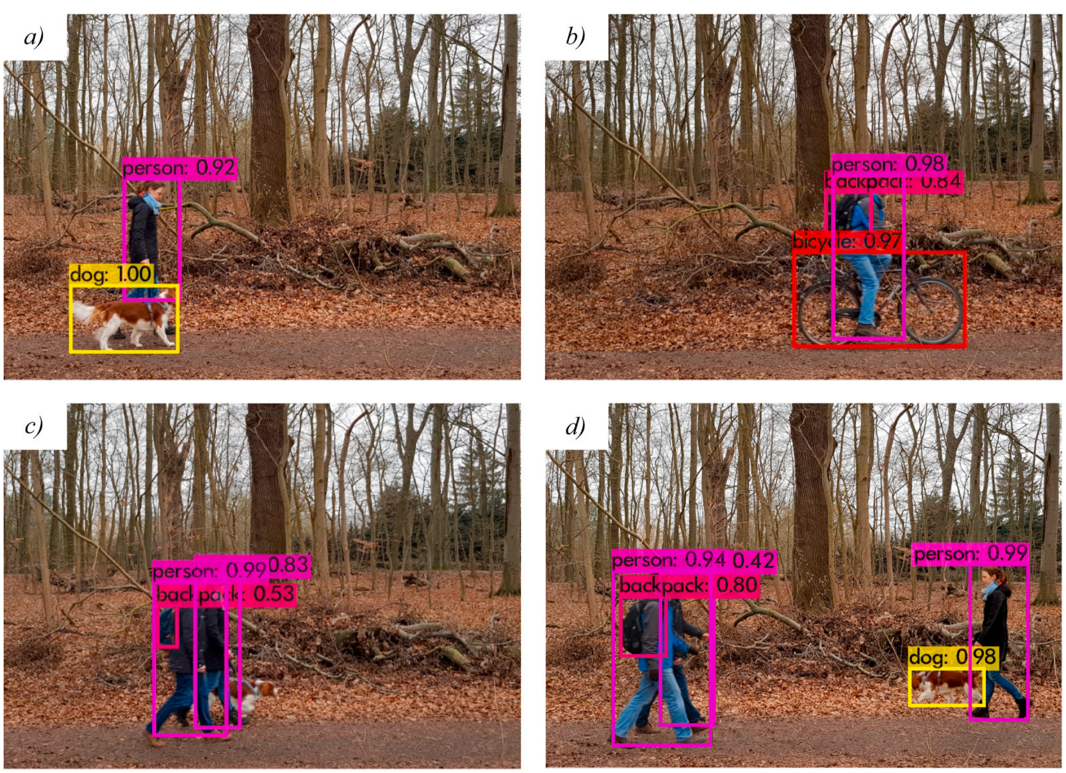

Fig. 5. Staged images with YOLO predictions illustrating its pre-trained semantic. Pink boxes refer to persons, red to equipment, and yellow to dogs. Each bounding box also includes a label specifying the predicted class and its certainty as decimal number. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4.2. Comparison of counting approaches

Based on the hourly observations, Table 7 shows the correlations between the different counting approaches for all concerned entrances. At entrance A, where manual in-situ observations were conducted next to a pressure sensor on five days, we found a strong and highly significant correlation ($R = 0.783$, $p < 0.001$). The respective linear model fitted through the 50 h of simultaneous observations further revealed that the automated approach does account for 88.4% of the visits counted by the manual in-situ observer (adj. $R^2 = 0.799$, $p < 0.001$).

Regarding the other two entrances where manual in-situ counting was conducted along with ongoing camera observations, only the manual camera evaluations at entrance B correlate significantly with the corresponding 44 h of manual in-situ observations. However, the regression models comparing the two camera evaluation approaches against manual in-situ observations are significant at both entrances. Unfortunately, at entrance D we could not relate the camera count to the in-situ counting, as the camera was stolen before the field observations started.

When comparing the camera evaluations against each other, both image evaluation approaches strongly correlate at very high significance levels ($Ør = 0.818$, $p < 0.001$). Further, it is interesting to note the highly significant regression coefficients between the two different image evaluation approaches as well. In respect to the differences of *Semantic 3* and YOLO, the slopes ranging between 0.478 and 0.570 can be interpreted as a percentage of Ø51.4% visitors entering the EFNR at these particular entrances.

To aid the interpretation of the partly contradictory results at entrance B, it is important to assess them specifically. When comparing the in-situ counts with the manual camera evaluations, similar numbers of visits were observed (Fig. 6a). However, further investigations of the scatterplots at entrance B (Fig. 6b and c) revealed four exceptional outliers as well. As documented in the manual image interpretation notes, on two days, several school classes entered EFNR before 09:00. When the group left the EFNR a few hours later, this was recorded by YOLO as its pre-trained semantic is not capable of differentiating directions (Fig. 6b). When the outliers highlighted rectangularly are excluded, the statistical measures parallel those of entrance C (see values in brackets, Table 7).

Further, $\text{Camera}^{\text{YOLO}}$ results also often correlate significantly but with low to medium strength to the in-situ personal counts and the pressure sensor at the other entrances (see Appendix 1). This shows that YOLO is able to reflect the visitation trends over the year even though observations did not always take place exactly at the same locations.

4.3. Annual visitation numbers to EFNR

Last but not least, the annual number of visits to ENFR was extrapolated for each of the counting approaches. We added a supplement to the manual in-situ counts to compensate for the missing observation hours (see Tables 3c and 4c). As YOLO was not able to distinguish walking directions, we rated its number proportional to the manual image assessment at entrances B, C and D (see regression coefficient Table 7). Fig. 7 depicts the number of visits for each entrance using the available visitor monitoring approach. Cumulated over the year 2015, the number of visits to the EFNR via entrances A-G ranges between 65,605 and 76,719.

5. Discussion

To assess the source of observational uncertainties, the significant differences between the visitor monitoring approaches need to be discussed. We do so in a chronological order. As deployed in multiple studies before (Job et al. 2005, Job, Merlin, et al., 2016; Job, 2008; Mayer et al., 2010), manual in-situ observations are a trusted approach to monitor visitors in protected areas. In this study, however, not all students instructed to conduct the observations showed up at the census line and recorded the counting on an hourly basis, although guidelines were provided. Hence, this approach does not only require a lot of resources in terms of interviewer organization, availability, training and payments but also control and data digitization. Additionally, this approach is not applicable on a daily basis for long-term projects and relies on a solid sampling scheme and extrapolation methods. In our study, a small sample size together with a high variance led to partly insignificant extrapolation coefficients and relatively large standard errors (see Table 2). Further, additional extrapolation is needed to cover visitation in the early morning, late evening and night hours. In this context, machine learning could provide interesting models for extrapolating visitor numbers (e.g. Rasanen et al., 2009; Taylor & Letham, 2017).

From a strictly statistical point of view, the pressure sensor may be considered a reliable instrument for visitor monitoring. In our explorative setup, this approach counted visitors all year without failures. Nevertheless, in some cases, we observed visitors avoiding stepping on the sensor, particularly on rainy days when there was some waterlogging on the slab. Therefore, careful choice of location and installation of pressure sensor slabs is crucial. Further, the sensitivity of the counter and the possibility of counts being triggered by baby strollers, dogs or wild animals has not been investigated so far. This raises some doubts about the precision of pressure sensor measurements (see also Ankre et al., 2016).

Before we discuss the image evaluation methods in particular, it is important to stress the limitations of triggered trail cameras in general. These relate not only to the legal and ethical aspects mentioned earlier, but also to maintenance, vandalism, and theft. Actually, all cameras were stolen before the end of our monitoring phase, which lead to the missing comparisons indicated in Table 7. Additionally, there were multiple gaps of missing observation dates, as cameras ran out of battery

Table 7
Statistical covariance analyses between counting approaches at entrances A-D. N corresponds to total hours of common observations.

	X	Y	N	Correlation	Regression	
				Pearson's R	Slope	Adjusted R^2
A	Manual in-situ	Pressure sensor	50	0.783 ***	0.884***	0.799
B	Manual in-situ	$\text{Camera}^{\text{Manual}}$	44	0.549 ***	0.791 ***	0.639
B	Manual in-situ	$\text{Camera}^{\text{YOLO}}$	44 (43)	0.007 (0.347 *)	1.317 *** (1.317 ***)	0.235 (0.575)
B	$\text{Camera}^{\text{YOLO}}$	$\text{Camera}^{\text{Manual}}$	1526 (1522)	0.786 *** (0.852 ***)	0.570 *** (0.639 ***)	0.715 (0.839)
C	Manual in-situ	$\text{Camera}^{\text{Manual}}$	17	0.412	0.523 ***	0.684
C	Manual in-situ	$\text{Camera}^{\text{YOLO}}$	17	0.335	0.979 ***	0.670
C	$\text{Camera}^{\text{YOLO}}$	$\text{Camera}^{\text{Manual}}$	691	0.847 ***	0.478 ***	0.846
D	Manual in-situ	$\text{Camera}^{\text{Manual}}$	0	/	/	/
D	Manual in-situ	$\text{Camera}^{\text{YOLO}}$	0	/	/	/
D	$\text{Camera}^{\text{YOLO}}$	$\text{Camera}^{\text{Manual}}$	379	0.821 ***	0.494 ***	0.825

Missing data is indicated by '/'. Regarding entrance B, coefficients after excluding four outliers with more than 100 visitors per hour are presented in brackets. See also Appendix 1 combining all approaches at all entrances; Levels of significance: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

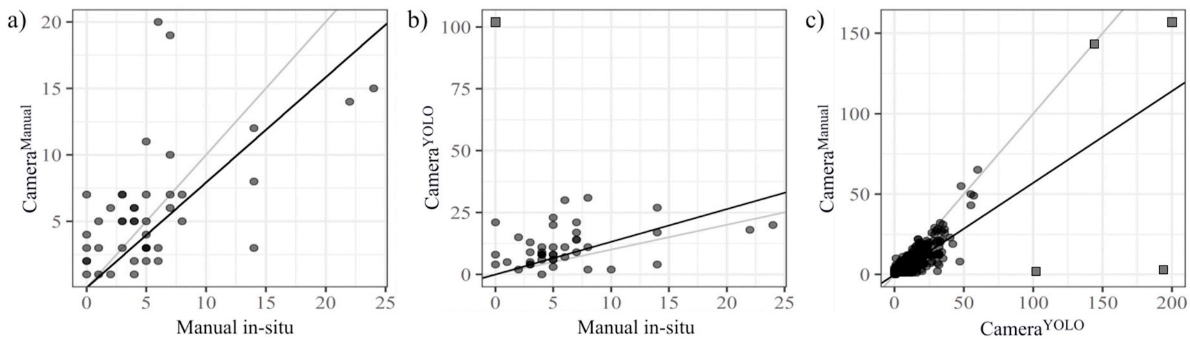

Fig. 6. Visualization of common observation data at entrance B, corresponding to Table 7. Squares are considered outliers and may be excluded.

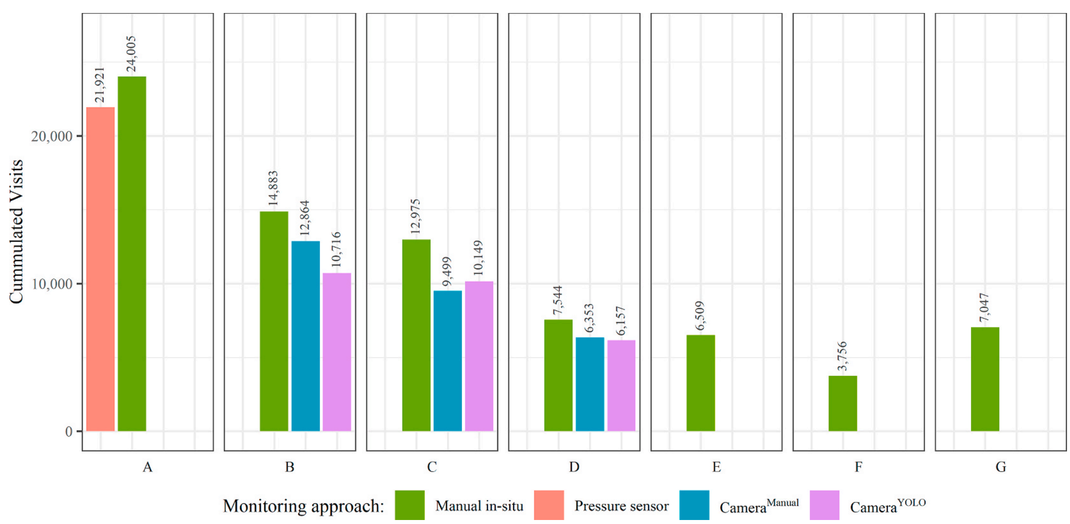

Fig. 7. Presentation of extrapolated annual number of visits per entrance and visitor monitoring approach.

power. Each triggered image captured by the device costs energy and, as a consequence, the battery runtime is proportional to the visitor number and can only be guessed. Also, regarding triggered trail cameras in particular, if a visitor moves faster than the device's activation time, the actual image will be blank (Miller et al., 2017). We assume this led to an undercount particularly at entrances B and C, where proportionally more cyclists were captured (see Table 5). Practitioners should consider the optimized camera protocol developed by Miller et al. (2017) in advance of starting future projects, particularly with regard to the camera's field of view, the trigger mechanism, and the file writing speed.

Nevertheless, cameras for visitor monitoring can balance well between the high information gain retrieved by in-situ observations and the corresponding costs. The manual image interpretation produces the most comprehensive visitation numbers. The records can easily be taken, and it offers the possibility of going back and forth between a set of images, allowing for very precise counting even in highly crowded scenes (Arnberger et al., 2005). In this context, it is also important to stress that if the manual image interpretation is done by different persons, clear semantics are required as a guideline to exclude uncertainties.

The pre-trained network by Redmon and Farhadi (2018) turned out to be transferable to visitor monitoring in protected areas and other recreational landscapes. The highly significant and strong correlations between Camera^{YOLO} and Camera^{Manual} show that this CNN is capable of producing comparable results in general; but, at dramatically reduced evaluation costs. Even without specialized hardware the images can be evaluated in the background of a common laptop computer. Further speed and accuracy enhancements may be expected from the updated YOLO v4 (Bochkovskiy et al., 2020). Additionally, we recommend testing and retraining other CNNs to meet the needs of visitor monitoring in protected areas. A semantically labeled training dataset would allow transferring the approach from detecting recreational equipment to actual outdoor recreation user groups. In this context, the training dataset could also include outdoor specific settings. This involves crowded images as well as fast moving and therefore blurred objects in particular. Finally, the automated image evaluation has another advantage. As analogous integrated circuits have become smaller and more effective in recent years, single-board computers may be deployed into the field (de Oliveira & Wehrmeister, 2018). Therewith, it is possible to evaluate the images in real time. While this live-processing approach has disadvantages in terms of auditability, as no imagery is actually recorded, it addresses the aforementioned privacy issues. That said, further configuration of the utilized software and hardware is needed to reach this goal. Regarding the possibilities of such a framework in general though, computer vision-based visitor monitoring is not only possible and suitable for outdoor recreation in protected areas but also for a much wider array of visitor counting settings. For example, Brown et al. (2016) gathered a network of webcams in cities or touristic sites to monitor the natural environments' phenology, which could be directly analyzed by YOLO or similar approaches too.

With respect to the annual extrapolation method, the higher annual number of visits from the manual in-situ counts might reflect a certain bias of actual counting days towards the spring season. This corresponds with the highest overall visitation frequency (see Fig. 3) leading to higher annual numbers due to the extrapolation compared to the automated approaches. The same reason might also explain why Table 2 reports higher differences between weekends/public holidays and weekdays (same weather conditions prevailing) for the manual in-situ counting than Table 3 for the pressure sensor. If, for instance, the counting days fall on public holidays in spring showing high frequentation (see Fig. 3), this leads to more pronounced differences than the pressure sensor values balanced over the whole year. Further, the lower difference of pressure sensor counts between weekends and weekdays with good weather compared to the manual in-situ counting could be also explained by different recreational behavior over the whole year. In midsummer, people most likely prefer the sea/beaches of the Baltic Sea farer away from the city for trips on weekends (also a possible explanation for lower overall visitation during midsummer, in combination with summer vacations spent elsewhere), while on weekdays they stay closer to their home and recreate in the EFNR. This could explain why the differences are lower when one considers the whole year as the weekend offers more outdoor recreation possibilities than weekdays. Likewise, the result from manual in-situ counting unexpectedly showed that on average slightly less visits occur to EFNR during good weather conditions compared to bad (Table 2). This appears counter intuitive at first, but it makes sense considering the recreational behavior of the citizens in Greifswald. EFNR is near and a very suitable destination for outdoor recreation even under adverse weather conditions, while trips to the Baltic Sea or other nearby destinations (e.g. the islands of Rügen or Usedom) are most likely preferred on warmer and sunnier days. This is also reflected in the camera-counting results for entrance B.

6. Conclusion

In this study we found that with triggered trail camera images and machine learning based computer vision the visitation frequency can be derived with high accuracy and comparatively low costs. This conclusion is based on conducting four visitor monitoring approaches at seven entrances to a protected forest area. The functional and physical differences of each entrance became evident with respect to overall frequentation, the fraction of visits outside the manual in-situ counting time window, and visitor types. The deployed counting approach needs to match these properties and the overall research questions. Manual in-situ counting can be very precise as it can distinguish between user groups, and also allows for parallel interviews which are required for a comprehensive management of protected areas (Spenceley et al., 2021). At the same time, this approach is very expensive for long-term monitoring while reducing the number of observation days introduces large uncertainties and compromises annual representativeness. Therefore, we recommend using this approach only at relatively high frequented sites where as many interviews as possible could be done in parallel to visitor counting. However, to further validate results, to better test CNNs at crowded locations and to add detailed yearly/seasonal visitation trends to visitor monitoring projects, we strongly argue for installing automatic counters (pressure or infrared sensors) and cameras at the exact same locations where manual counting is done to ensure the highest data comparability. Analog thereto an investigation of automated cameras' accuracy along different constraints such as crowding and weather would be interesting.

Less or very low frequented sites should be covered by automatic counters (or cameras). Pressure sensors are technically reliable and suitable for long-term monitoring projects. Although such devices are relatively expensive to buy, not easily portable and their installation requires time and experience, there were no issues related to maintenance or vandalism. However, there are some limitations with respect to the broad requirements of managing protected areas. For instance, they cannot distinguish between user groups and some doubts remain about the specificity of the pressure sensor to count humans only.

The use of automated cameras, however, allows for comprehensive visitor monitoring with regard to reproducibility and differentiating user groups (which cannot be done by conventional automatic counters) – an important aspect in mitigating recreational frictions or wildlife disturbance conflicts. However, triggered trail cameras, more commonly used to survey wildlife, come with the shortcomings of huge time and financial costs when evaluating data, together with the risk of human error during image processing. The use of computer vision can drastically reduce these problems. The field experiments conducted in this study showed that YOLO (Redmon & Farhadi, 2018), as one of many CNNs, reliably derived similar visitation counts compared to manual image evaluations. Nevertheless, camera installation takes time and effort, regular maintenance (batteries, storage cards) and the automated evaluation requires specific hardware and expertise as well. Other non-trivial issues are theft, vandalism and the short lifespan of batteries, especially in the winter season where potential snow cover leads to impairment of camera functions. Further, legal issues must be clarified whenever the use of cameras is planned. Hopefully, future works will spawn live-processing devices, so that no imagery needs to be stored at all. This however, would require a very reliable algorithm. We therefore strongly recommend retraining a CNN particular to the needs of visitor monitoring in protected areas and other recreational landscapes. In doing so, automated computer vision will detect a manifold width of detectable objects, like semantic visitor types, vehicle types (as suggested by Thórhallsdóttir, Ólafsson, & Jóhannesson (2021)) and much more outdoor equipment in the future. In case of wildlife monitoring, the *Snapshot Serengeti* project (Swanson et al., 2015) challenged big data scientists to build the best algorithms. So, why not launch a comparable initiative for the issues mentioned regarding visitation in outdoor-recreation settings? Along thereto, CNNs could also be deployed in the emerging field of social media-based visitor monitoring (Ghermandi & Sinclair, 2019; Sinclair et al., 2020a, 2020b; Teles de la Mota & Pickering, 2020) where it might help to automatically detect and analyze the contents of pictures taken and posted by visitors. Lastly, using the very same camera sensor, multiple aspects of the respective protected areas and other recreational landscapes can be monitored in parallel: wildlife (e.g. Falzon et al., 2020), vegetation (e.g. Brown et al., 2016) and visitors too.

Ethical disclosure

All data was collected and handled regarding the German data privacy regulations. Further the persons shown in Fig. 5 agree on their imagery being published and are credited in the acknowledgments. No animals were hurt.

CRediT authorship contribution statement

Jeroen Staab: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Erica Udas:** Conceptualization, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing. **Marius Mayer:** Conceptualization, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration. **Hannes Taubenböck:** Resources, Writing – review & editing, Supervision. **Hubert Job:** Resources, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Wilhelm Steingrube and Norman Volkmann for their important help during the data collection pertaining to this study. Further, we thank Samuel Thomas for language editing the manuscript, Winfried Weber for the cartographic design of Fig. 1, as well as Julia Sinner for illustrating Fig. 2 and Mareike Garms, her dog Anton and Philipp Sacher for staging the exemplary scenes shown in Fig. 5. Last but not least we thank both reviewers for their constructive feedback.

Appendix 1. Statistical covariance analyses between counting approaches combining all approaches at all entrances. Comparisons are presented into one direction only. They can be reversed by applying 1/slope

X		Y		Correlation Pearson's R	Regression Slope	Adjusted R^2
Entrance	Method	Entrance	Method			
A	InSitu	A	Ecocounter	0.783 ***	0.884 ***	0.799
A	InSitu	B	InSitu	0.403 **	0.435 ***	0.498
A	InSitu	B	$Camera^{Manual}$	0.727 ***	0.566 ***	0.792
A	InSitu	B	$Camera^{YOLO}$	0.131	0.885 ***	0.274
A	InSitu	C	InSitu	0.338 *	0.285 ***	0.457
A	InSitu	C	$Camera^{Manual}$	0.539 *	0.214 ***	0.715
A	InSitu	C	$Camera^{YOLO}$	0.200	0.366 ***	0.575
A	InSitu	D	InSitu	0.356 *	0.230 ***	0.456
A	InSitu	E	InSitu	0.357	0.281 **	0.488
A	InSitu	F	InSitu	0.313	0.077 **	0.517
A	InSitu	G	InSitu	−0.212	0.077	−0.044
A	Ecocounter	B	InSitu	0.267 **	0.401 ***	0.417
A	Ecocounter	B	$Camera^{Manual}$	0.304 ***	0.553 ***	0.279
A	Ecocounter	B	$Camera^{YOLO}$	0.274 ***	0.714 ***	0.254
A	Ecocounter	C	InSitu	0.314 ***	0.325 ***	0.424
A	Ecocounter	C	$Camera^{Manual}$	0.447 ***	0.267 ***	0.477
A	Ecocounter	C	$Camera^{YOLO}$	0.489 ***	0.514 ***	0.488
A	Ecocounter	D	InSitu	0.398 ***	0.199 ***	0.417
A	Ecocounter	D	$Camera^{Manual}$	0.380 ***	0.184 ***	0.463
A	Ecocounter	D	$Camera^{YOLO}$	0.516 ***	0.336 ***	0.528
A	Ecocounter	E	InSitu	0.429 ***	0.184 ***	0.461
A	Ecocounter	F	InSitu	0.267 ***	0.114 ***	0.328
A	Ecocounter	G	InSitu	−0.089 ***	0.166#	0.068
B	InSitu	B	$Camera^{Manual}$	0.549 ***	0.791 ***	0.639
B	InSitu	B	$Camera^{YOLO}$	0.007	1.317 ***	0.235
B	InSitu	C	InSitu	0.157	0.425 ***	0.325
B	InSitu	C	$Camera^{Manual}$	0.283	0.303 ***	0.592
B	InSitu	C	$Camera^{YOLO}$	0.301	0.585 ***	0.620
B	InSitu	D	InSitu	0.245 *	0.285 ***	0.316
B	InSitu	E	InSitu	0.162	0.236 ***	0.298
B	InSitu	F	InSitu	0.250	0.143 ***	0.304
B	InSitu	G	InSitu	−0.072	0.220 *	0.077
B	$Camera^{Manual}$	B	$Camera^{YOLO}$	0.786 ***	1.255 ***	0.715
B	$Camera^{Manual}$	C	InSitu	0.344 *	0.521 ***	0.493
B	$Camera^{Manual}$	C	$Camera^{Manual}$	0.049	0.123 ***	0.124
B	$Camera^{Manual}$	C	$Camera^{YOLO}$	0.100 *	0.272 ***	0.141
B	$Camera^{Manual}$	D	InSitu	0.385 *	0.367 ***	0.430
B	$Camera^{Manual}$	D	$Camera^{Manual}$	0.151 *	0.278 ***	0.256
B	$Camera^{Manual}$	D	$Camera^{YOLO}$	0.189 **	0.439 ***	0.234
B	$Camera^{Manual}$	E	InSitu	0.219	0.367 ***	0.360
B	$Camera^{Manual}$	F	InSitu	0.153	0.150 **	0.320
B	$Camera^{Manual}$	G	InSitu	−0.287	0.169	−0.003
B	$Camera^{YOLO}$	C	InSitu	−0.014	0.136 ***	0.208
B	$Camera^{YOLO}$	C	$Camera^{Manual}$	0.048	0.085 ***	0.126
B	$Camera^{YOLO}$	C	$Camera^{YOLO}$	0.095 *	0.184 ***	0.140
B	$Camera^{YOLO}$	D	InSitu	0.388 *	0.236 ***	0.466
B	$Camera^{YOLO}$	D	$Camera^{Manual}$	0.102	0.206 ***	0.216
B	$Camera^{YOLO}$	D	$Camera^{YOLO}$	0.127 *	0.304 ***	0.185
B	$Camera^{YOLO}$	E	InSitu	0.418 #	0.265 ***	0.552
B	$Camera^{YOLO}$	F	InSitu	0.264	0.098 ***	0.399
B	$Camera^{YOLO}$	G	InSitu	−0.184	0.034	−0.028
C	InSitu	C	$Camera^{Manual}$	0.412	0.523 ***	0.684
C	InSitu	C	$Camera^{YOLO}$	0.335	0.979 ***	0.670
C	InSitu	D	InSitu	0.127	0.297 ***	0.239
C	InSitu	E	InSitu	0.169	0.269 ***	0.253
C	InSitu	F	InSitu	0.209	0.196 ***	0.273
C	InSitu	G	InSitu	−0.098	0.438 *	0.096
C	$Camera^{Manual}$	C	$Camera^{YOLO}$	0.847 ***	1.772 ***	0.846
C	$Camera^{Manual}$	D	InSitu	0.154	1.124 **	0.507
C	$Camera^{Manual}$	D	$Camera^{Manual}$	0.310 ***	0.494 ***	0.419
C	$Camera^{Manual}$	D	$Camera^{YOLO}$	0.444 ***	0.934 ***	0.493
C	$Camera^{Manual}$	E	InSitu	−0.544	0.889	0.369
C	$Camera^{Manual}$	F	InSitu	0.322	0.371 ***	0.694
C	$Camera^{Manual}$	G	InSitu	0.234	0.390 *	0.300
C	$Camera^{YOLO}$	D	InSitu	−0.007	0.572 **	0.473
C	$Camera^{YOLO}$	D	$Camera^{Manual}$	0.397 ***	0.266 ***	0.456
C	$Camera^{YOLO}$	D	$Camera^{YOLO}$	0.497 ***	0.485 ***	0.520
C	$Camera^{YOLO}$	E	InSitu	0.570	0.356 *	0.774
C	$Camera^{YOLO}$	F	InSitu	0.252	0.204 **	0.663

(*continued on next page*)

(*continued*)

X		Y		Correlation Pearson's R	Regression Slope	Adjusted R^2
Entrance	Method	Entrance	Method			
C	$Camera^{YOLO}$	G	InSitu	0.094	0.171 *	0.253
D	InSitu	E	InSitu	0.179	0.524 ***	0.234
D	InSitu	F	InSitu	0.314	0.185 **	0.301
D	InSitu	G	InSitu	0.064	0.554 **	0.242
D	$Camera^{Manual}$	D	$Camera^{YOLO}$	0.821 ***	1.669 ***	0.825
E	InSitu	F	InSitu	−0.115	0.155	0.052
E	InSitu	G	InSitu	0.028	0.626 **	0.225
F	InSitu	G	InSitu	−0.033	1.045 *	0.186

Levels of significance: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, #$p < 0.1$.

References

Ankre, R., Fredman, P., & Lindhagen, A. (2016). Managers' experiences of visitor monitoring in Swedish outdoor recreational areas. *Journal of Outdoor Recreation and Tourism, 14*, 35–40. https://doi.org/10.1016/j.jort.2016.04.008

Arnberger, A. (2007). Internationale Entwicklungen im Besuchermonitoring – Ein Überblick. In Biosphärenreservat Vessertal-Thüringer Wald. In Biosphärenreservat Vessertal-Thüringer Wald (Ed.), *Besuchermonitoring und ökonomische Effekte in Nationalen Naturlandschaften* (pp. 8–17). Schmiedefeld: Biosphärenreservat Vessertal-Thüringer Wald.

Arnberger, A., Eder, R., & Preisel, H. (2016). Tagestourismus oder Wohnumfeldnutzung? Ein Vergleich der Besuchsintensitäten und -muster von drei Erholungs- und Schutzgebieten in Wien. *Zeitschrift für Tourismuswissenschaft, 8*(2), 199–221. https://doi.org/10.1515/tw-2016-0018

Arnberger, A., Haider, W., & Brandenburg, C. (2005). Evaluating visitor-monitoring techniques: A comparison of counting and video observation data. *Environmental Management, 36*(2), 317–327. https://doi.org/10.1007/s00267-004-8201-6

Balmford, A., Green, J. M. H., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M., & Manica, A. (2015). Walk on the wild side: Estimating the global magnitude of visits to protected areas. *PLoS Biology, 13*(2), 1–6. https://doi.org/10.1371/journal.pbio.1002074

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. ArXiv:2004.10934 http://arxiv.org/abs/2004.10934.

Brown, T. B., Hultine, K. R., Steltzer, H., Denny, E. G., Denslow, M. W., Granados, J., Henderson, S., Moore, D., Nagai, S., San Clements, M., Sánchez-Azofeifa, A., Sonnentag, O., Tazik, D., & Richardson, A. D. (2016). Using phenocams to monitor our changing Earth: Toward a global phenocam network. *Frontiers in Ecology and the Environment, 14*(2), 84–93. https://doi.org/10.1002/fee.1222

Brunetti, A., Buongiorno, D., Trotta, G. F., & Bevilacqua, V. (2018). Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing, 300*, 17–33. https://doi.org/10.1016/j.neucom.2018.01.092

Bundesdatenschutzgesetz (Bdsg). (2018). *German federal data protection act.* https://www.gesetze-im-internet.de/bdsg_2018/. (Accessed 1 April 2020) Accessed.

Cessford, G. R., & Burns, R. (2008). *Monitoring visitor numbers in New Zealand national parks and protected areas: A literature review and development summary.* Wellington, N. Z.: Dept of Conservation.

Cessford, G. R., & Muhar, A. (2003). Monitoring options for visitor numbers in national parks and natural areas. *Journal for Nature Conservation, 11*(4), 240–250. https://doi.org/10.1078/1617-1381-00055

Czachs, C., & Brandenburg, C. (2014). Visitor monitoring with time lapse trail cameras. In M. Reimann, K. Sepp, E. Pärna, & R. Tuula (Eds.), *Management for protection and sustainable development. Proceedings of the 7th international conference on monitoring and management of visitors in recreational and protected areas* (pp. 303–305). Tallinn, Estonia: Tallinn University.

Dollár, P., Wojek, C., Schiele, B., & Perona, P. (2012). Pedestrian detection: An evaluation of the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*(4), 743–761. https://doi.org/10.1109/TPAMI.2011.155

Falzon, G., Lawson, C., Cheung, K.-W., Vernes, K., Ballard, G. A., Fleming, P. J. S., Glen, A. S., Milne, H., Mather-Zardain, A., & Meek, P. D. (2020). *ClassifyMe*: A field-scouting software for the identification of wildlife in camera trap images. *Animals, 10*(1), 1–16. https://doi.org/10.3390/ani10010058

Ghermandi, A., & Sinclair, M. (2019). Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change, 55*, 36–47. https://doi.org/10.1016/j.gloenvcha.2019.02.003

Gomez Villa, A., Salazar, A., & Vargas, F. (2017). Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecological Informatics, 41*, 24–32. https://doi.org/10.1016/j.ecoinf.2017.07.004

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing, 187*, 27–48. https://doi.org/10.1016/j.neucom.2015.09.116

Hannemann, T., & Job, H. (2003). Destination "Deutsche Nationalparke" als touristische Marke. *Tourism Review, 58*(2), 6–17. https://doi.org/10.1108/eb058404

Han, F., Yao, J., Zhu, H., & Wang, C. (2020). Underwater image processing and object detection based on deep CNN method. *Journal of Sensors, 2020*, Article e6707328. https://doi.org/10.1155/2020/6707328

Highfill, T., & Franks, C. (2019). Measuring the U.S. outdoor recreation economy, 2012–2016. *Journal of Outdoor Recreation and Tourism, 27*, 100233. https://doi.org/10.1016/j.jort.2019.100233

Hodges, L. (2009). Ultrasonic and Passive Infrared Sensor Integration for Dual Technology User Detection Sensors. http://www.egr.msu.edu/classes/ece480/capstone/fall09/group05/docs/ece480_dt5_application_note_lhodges.pdf. (Accessed 1 April 2020) Accessed.

Hornback, K. E., & Eagles, P. F. J. (1999). *Guidelines for public use measurement and reporting at parks and protected areas.* Gland: Cambridge.

Job, H. (2008). Estimating the regional economic impact of tourism to national parks: Two case studies from Germany. *GAIA - Ecological Perspectives for Science and Society, 17*(S1), 134–142. https://doi.org/10.14512/gaia.17.S1.11

Job, H., Harrer, B., Metzler, D., & Hajizadeh-Alamdary, D. (2005). *Ökonomische Effekte von Großschutzgebieten.* Bonn-Bad Godesberg: Bundesamt für Naturschutz.

Job, H., Merlin, C., Metzler, D., Schamel, J., & Woltering, M. (2016a). *Regionalwirtschaftliche Effekte durch Naturtourismus in deutschen Nationalparken als Beitrag zum Integrativen Monitoring-Programm für Großschutzgebiete.* Bonn-Bad Godesberg: Bundesamt für Naturschutz.

Job, H., Schamel, J., & Butzmann, E. (2016b). Besuchermanagement in Großschutzgebieten im Zeitalter moderner Informations- und Kommunikationstechnologien. *Natur und Landschaft, 91*(1), 32–38.

Kahler, A., & Arnberger, A. (2008). A comparison of passive infrared counter results with time lapse video monitoring at a shared urban recreational trail. In A. Raschi, & S. Trampetti (Eds.), *Management for protection and sustainable development* (pp. 485–489). Italy: Montecatini Terme.

Kajala, L., Almik, A., Dahl, R., Diksaite, L., Erkkonen, J., Fredman, P., Jensen, F., Sondergaard, F., Karoles, K., Sievänen, T., Sov-Petersen, H., Vistad, O., & Wallsten, P. (2007). *Visitor monitoring in nature areas. A manual based on experiences from the Nordic and Baltic countries.* Stockholm, Sweden: The Swedish Environmental Protection Agency.

Kays, R., Tilak, S., Kranstauber, B., Jansen, P. A., Carbone, C., Rowcliffe, M., & He, Z. (2011). Camera traps as sensor networks for monitoring animal communities. *International Journal of Research and Reviews in Wireless Sensor Networks, 1*(2), 19–29.

Lupp, G., Kantelberg, V., Förster, B., Naumann, J., Honert, C., Markmann, T., Koch, M., Schreiber, R., & Pauleit, S. (2016). Vorsicht Kamera! Besuchermonitoring mit Wildkameras. *LWF Aktuell, 111*, 14–16.

Mayer, M., Wasem, K., Gehring, K., Pütz, M., Roschewitz, A., & Siegrist, D. (2009). *Wirtschaftliche Bedeutung des naturnahen Tourismus im Simmental und Diemtigtal – Regionalökonomische Effekte und Erfolgsfaktoren.* Birmensdorf/Rapperswil: Eidg. Forschungsanstalt für Wald, Schnee und Landschaft WSL/Hochschule für Technik Rapperswil HSR.

Mayer, M., Müller, M., Woltering, M., Arnegger, J., & Job, H. (2010). The economic impact of tourism in six German national parks. *Landscape and Urban Planning, 97*(2), 73–82. https://doi.org/10.1016/j.landurbplan.2010.04.013

Mayer, M., & Woltering, M. (2018). Assessing and valuing the recreational ecosystem services of Germany's national parks using travel cost models. *Ecosystem Services, 31* (Part C), 371–386. https://doi.org/10.1016/j.ecoser.2017.12.009

Miller, A. B., Leung, Y.-F., & Kays, R. (2017). Coupling visitor and wildlife monitoring in protected areas using camera traps. *Journal of Outdoor Recreation and Tourism, 17*, 44–53. https://doi.org/10.1016/j.jort.2016.09.007

Millhäusler, A., Anderwald, A., Haeni, M., & Haller, R. M. (2016). Publicity, economics and weather – changes in visitor numbers to a European National Park over 8 years. *Journal of Outdoor Recreation and Tourism, 16*, 50–57. https://doi.org/10.1016/j.jort.2016.09.005

Muhar, A., Arnberger, A., & Brandenburg, C. (2002). Methods for visitor monitoring in recreational and protected areas: An overview. In A. Muhar, A. Arnberger, & C. Brandenburg (Eds.), *Monitoring and management of visitor flows in recreational and protected areas* (pp. 1–6). Wien: Universität für Bodenkultur.

Muhar, A., Zemann, R., & Lengauer, M. (1995). Permanent time-lapse video recording for the quantification of recreational activities. In *Proc. Decision support systems 2001. Resource Technology 94* (pp. 219–229). Bethesda: Am. Soc. Photogrammetry and Remote Sensing.

de Oliveira, D., & Wehrmeister, M. (2018). Using deep learning and low-cost RGB and thermal cameras to detect pedestrians in aerial images captured by multirotor UAV. *Sensors, 18*(7), 2244. https://doi.org/10.3390/s18072244

Pettebone, D., Newman, P., & Lawson, S. R. (2010). Estimating visitor use at attraction sites and trailheads in Yosemite National Park using automated visitor counters.

Landscape and Urban Planning, 97(4), 229–238. https://doi.org/10.1016/j.landurbplan.2010.06.006
Pickering, C., Dario Rossi, S., Hernando, A., & Barros, A. (2018). Current knowledge and future research directions for the monitoring and management of visitors in recreational and protected areas. *Journal of Outdoor Recreation and Tourism, 21*, 10–18. https://doi.org/10.1016/j.jort.2017.11.002
Rasanen, T., Niska, H., Hiltunen, T., Tiirikainen, J., & Kolehmainen, M. (2009). Predictive system for monitoring regional visitor attendance levels in large recreational areas. *Journal of Environmental Informatics, 13*(1), 45–55. https://doi.org/10.3808/jei.200900139
Rathmann, J., Beck, C., Flutur, S., Seiderer, A., Aslan, I., & André, E. (2020). Towards quantifying Forest Recreation: Exploring outdoor thermal physiology and human well-being along exemplary pathways in a central European urban forest (Augsburg, SE-Germany). *Urban Forestry and Urban Greening, 49*, 126622. https://doi.org/10.1016/j.ufug.2020.126622
Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only Look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016* (pp. 779–788). The Computer Vision Foundation.
Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017* (pp. 7263–7271). The Computer Vision Foundation.
Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement*. arXiv: 1804.02767 [cs.CV].
Rein, H., & Baláš, M. (2015). *Die Wertschöpfung des Tourismus im Nationalpark Unteres Odertal. Vergleichsstudie 2007/08–2013/14*. Eberswalde: Hochschule für Nachhaltige Entwicklung.
Rein, H., Roberts, D., & Baláš, M. (2019). *Bedeutung des Nationalparks für die touristische Entwicklung der Welterberegion Wartburg Hainich. Ergebnisse des sozioökonomischen Monitorings. Nationalpark-Verwaltung Hainich: Bad Langensalza*.
Rupf, R., Haider, W., & Pröbstl, U. (2014). Hikers and mountain bikers–do they fight like cats and dogs? In M. Reimann, K. Sepp, E. Pärna, & R. Tuula (Eds.), *Management for protection and sustainable development. Proceedings of the 7th international conference on monitoring and management of visitors in recreational and protected areas* (pp. 253–255). Tallinn: Tallinn University.
Schägner, J. P., Brander, L., Paracchini, M. L., Maes, J., Gollnow, F., & Bertzky, B. (2018). Spatial dimensions of recreational ecosystem service values: A review of meta-analyses and a combination of meta-analytic value-transfer and GIS. *Ecosystem Services, 31*, 395–409. https://doi.org/10.1016/j.ecoser.2018.03.003
Schägner, J. P., Maes, J., Brander, L., Paracchini, M. L., Hartje, V., & Dubois, G. (2017). Monitoring recreation across European nature areas: A geo-database of visitor counts, a review of literature and a call for a visitor counting reporting standard. *Journal of Outdoor Recreation and Tourism, 18*, 44–55. https://doi.org/10.1016/j.jort.2017.02.004
Schamel, J., & Job, H. (2013). Crowding in Germany's national parks: The case of the low mountain range saxon Switzerland national park. *Eco. Mont – Journal on Protected Mountain Areas Research and Management, 5*(1), 27–34. https://doi.org/10.1553/eco.mont-5-1s27
Schneider, S., Taylor, G. W., & Kremer, S. (2018). Deep learning object detection methods for ecological camera trap data. In *2018 15th conference on computer and robot vision* (pp. 321–328). CRV. doi:10.1109/CRV.2018.00052.
Sinclair, M., Mayer, M., Woltering, M., & Ghermandi, A. (2020a). Using social media to estimate visitor provenance and patterns of recreation in Germany's national parks. *Journal of Environmental Management, 263*, 110418. https://doi.org/10.1016/j.jenvman.2020.110418
Sinclair, M., Mayer, M., Woltering, M., & Ghermandi, A. (2020b). Valuing nature-based recreation using a crowdsourced travel cost method: A comparison to onsite survey data and value transfer. *Ecosystem Services, 45*, 101165. https://doi.org/10.1016/j.ecoser.2020.101165
Spenceley, A., Schägner, J. P., Engels, B., Engelbauer, M., Erkkonen, J., Job, H., … Woltering, M. (2021). *Visitors count! Guidance for protected areas on the economic analysis of visitation*. Paris, Bonn, Ispra: UNESCO, BfN, EUJRC.
Stiller, D., Stark, T., Wurm, M., Dech, S., & Taubenböck, H. (2019). Large-scale building extraction in very high resolution aerial imagery using Mask R-CNN. In *IEEE-CPS joint urban remote sensing event (JURSE)*. Vannes, France.
Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A., & Packer, C. (2015). Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data, 2*(1), 1–14. https://doi.org/10.1038/sdata.2015.26
Taye, F. A., Abildtrup, J., Mayer, M., Ščasný, M., & Lundhede, T. (2019). Childhood experience in forest recreation practices: Evidence from nine European countries. *Urban Forestry and Urban Greening, 46*, 126471. https://doi.org/10.1016/j.ufug.2019.126471
Taylor, S. J., & Letham, B. (2017). Forecasting at scale. *PeerJ Preprints, 5*, Article e3190v2. https://doi.org/10.7287/peerj.preprints.3190v2
Teles de la Mota, V., & Pickering, C. (2020). Using social media to assess nature-based tourism: Current research and future trends. *Journal of Outdoor Recreation and Tourism, 30*, 100295. https://doi.org/10.1016/j.jort.2020.100295
Thórhallsdóttir, G., Ólafsson, R., & Jóhannesson, G. T. (2021). A methodology of estimating visitor numbers at an Icelandic destination using a vehicle counter and a radar. *Journal of Outdoor Recreation and Tourism, 35*, Article 100378. https://doi.org/10.1016/j.jort.2021.100378.
Udas, E., Wölk, M., & Wilmking, M. (2018). The "carbon-neutral university" - a study from Germany. *International Journal of Sustainability in Higher Education, 19*(1), 130–145. https://doi.org/10.1108/IJSHE-05-2016-0089
Woltering, M. (2012). *Tourismus und Regionalentwicklung in deutschen Nationalparken: Regionalwirtschaftliche Wirkungsanalyse des Tourismus als Schwerpunkt eines sozioökonomischen Monitoringsystems*. Würzburg: University of Würzburg Press.
Wurm, M., Stark, T., Zhu, X. X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing, 150*, 59–69. https://doi.org/10.1016/j.isprsjprs.2019.02.006
Yousif, H., Yuan, J., Kays, R., & He, Z. (2019). Animal scanner: Software for classifying humans, animals, and empty frames in camera trap images. *Ecology and Evolution, 9*(4), 1578–1589. https://doi.org/10.1002/ece3.4747
Zhang, S., Benenson, R., Omran, M., Hosang, J., & Schiele, B. (2017). Towards reaching human performance in pedestrian detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*(4), 973–986. https://doi.org/10.1109/TPAMI.2017.2700460
Ziesler, P. S., & Pettebone, D. (2018). Counting on visitors: A review of methods and applications for the national park service's visitor use statistics program. *Journal of Park and Recreation Administration, 36*, 39–55. https://doi.org/10.18666/JPRA-2018-V36-I1-8104

Jeroen Staab received the M.Sc. degree in Applied Human Geography from the Julius-Maximilians-University Würzburg, Germany in 2017. After that, he joined the German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Weβling, Germany, as a Research Associate. He is currently working towards the Ph.D. degree with the Humbold University Berlin, Germany. His data driven research interests involve computer vision, machine learning and complex geostatistical modelling contributing to regional development and informed policy making.

Erica Udas is a landscape ecologist and a forester pursuing her Ph.D. on tradeoffs and synergies between forest ecosystem services at University Greifswald, Germany. Currently, she is working as an Ecosystem Analyst at the International Center for Integrated Mountain Development (ICIMOD), Nepal. Her research interests cover assessment and valuation of ecosystem services, forest ecology, building socio-ecological resilience to impacts of climate change and other shocks, sustainable mountain development, and green recovery.

Marius Mayer is an economic geographer (Diploma from Ludwig-Maximilians University München, Germany in 2006) working as post-doctoral researcher at the Department of Strategic Management, Marketing and Tourism (Team SME and tourism) of University Innsbruck, Austria. He worked as junior professor for economic geography and tourism at University Greifswald, Germany from 2013 to 2020 and has a PhD (Dr. rer. nat.) in geography (2012) from Julius-Maximilians-University Würzburg, Germany. His research interests cover visitor monitoring and management, the economic impact and valuation of outdoor recreation inside and outside of protected areas, the economics of protected areas, climate change impacts on and adaptation of outdoor recreation as well as regional development.

Hannes Taubenböck received the Diploma in geography from the Ludwig-Maximilians University München, Germany, in 2004, and the Ph.D.degree (Dr.rer.nat.) in geography from the Julius Maximilian's University of Würzburg, Germany, in 2008. In 2005, he joined the German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Weβling, Germany. After a postdoctoral research phase with the University of Würzburg (2007–2010), he returned in 2010 to DLR–DFD as a Scientific Employee. In 2013, he became the Head of the "City and Society" team. In 2019, he habilitated at the University of Würzburg in Geography. His current research interests include urban remote sensing topics, from the development of algorithms for information extraction to value adding classification products for findings in urban geography.

Hubert Job was Associate Professor of Economic Geography at Ludwig-Maximilians-University München, Germany, from 2000 until 2008. Since then he is Full Professor of Geography and Regional Science at Julius-Maximilians-University Würzburg, Germany. He has conducted many research projects on protected areas and thus there are numerous publications related to park management issues.