Landscape and Urban Planning 219 (2022) 104315

Contents lists available at ScienceDirect

Landscape and Urban Planning

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

Promises and pitfalls of using computer vision to make inferences about landscape preferences: Evidence from an urban-proximate park system

Emily J. Wilkins [a,b,*], Derek Van Berkel [c], Hongchao Zhang [a,b], Monica A. Dorning [d], Scott M. Beck [e], Jordan W. Smith [a,b]

[a] Department of Environment and Society, Utah State University, Logan, UT 84322, USA
[b] Institute of Outdoor Recreation and Tourism, Utah State University, Logan, UT 84322, USA
[c] School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109, USA
[d] Department of Geosciences and Natural Resources, Western Carolina University, Cullowhee, NC 28723, USA
[e] American Farmland Trust, Washington, DC 20036, USA

HIGHLIGHTS

- Differences in landscape preferences identified from surveys and social media.
- Development was not preferred by visitors, but was often in photo content.
- No large differences between those who share images online and those who do not.
- Some differences in automated versus manual photograph content analysis.

ARTICLE INFO

Keywords:
Machine Learning
Social media
Landscape preferences
Image content analysis
Protected area
Visitor survey

ABSTRACT

The ubiquitous use of the internet and social media has provided social and spatial scientists with a wealth of data from which inferences about landscape preferences can be gained. These data are increasingly being used as an alternative to data collected from surveys of recreationists. While the rapidly growing body of research using social media is impressive, little work has been done to compare the image content of social media to preferences elucidated via more traditional methods. We compare the landscape features derived through a computer vision algorithm used to analyze social media photographs with preferences derived through a traditional on-site intercept survey. We found that landscape features identified through the computer vision algorithm were, by and large, significantly different compared to landscape features that park users said improved their recreational experiences. Additionally, we did not find substantial differences in landscape preferences between visitors who share photographs of their park visit on social media and those who do not. We suggest a diversity of data sources and analytical methods should be used in a complementary and comparative way. Our analysis here suggests both surveys and social media images can provide important insights about landscape preferences, but neither in isolation is perfect.

1. Introduction

Publicly accessible parks and open spaces provide a variety of psychological, physiological, and social benefits to the individuals who access them for recreation. However, the ability of recreationists to realize these benefits is heavily dependent upon the aesthetic characteristics and visual appeal of these spaces (Velarde, Fry, & Tveit, 2007; Wang, Zhao, Meitner, Hu, & Xu, 2019). Parklands and open spaces that provide opportunities for individuals to experience more natural settings and extensive views are preferred relative to landscapes without these characteristics (Kaplan, 1995). Strategic investments in outdoor recreation infrastructure, such as new trails, picnic areas, and campsites, which maximize views of landscape features that tend to be appealing to outdoor recreationists can subsequently increase not only the demand

* Corresponding author now at: U.S. Geological Survey, Fort Collins Science Center, 2150 Centre Ave, Bldg C., Fort Collins, CO 80526, USA.
E-mail addresses: ewilkins@usgs.gov (E.J. Wilkins), dbvanber@umich.edu (D. Van Berkel), hongchao.zhang@aggiemail.usu.edu (H. Zhang), jordan.smith@usu.edu (J.W. Smith).

https://doi.org/10.1016/j.landurbplan.2021.104315
Received 10 May 2021; Received in revised form 16 November 2021; Accepted 20 November 2021
Available online 2 December 2021

for that infrastructure but also the ability of visitors to obtain desired benefits (Rosenberger, Bergerson, & Kline, 2009; Stein & Lee, 1995).

Outdoor recreation planners, land use planners, and researchers often turn to surveys as the default method to understand the landscape preferences of visitors (e.g., DeLucio & Múgica, 1994; Howley, 2011; Manning, 2011; Múgica & De Lucio, 1996). Surveys are frequently administered on-site to ascertain current users' stated preferences for landscape features (e.g., views of waterbodies, mountains, forests, etc.) (e.g., Jiang & Yuan, 2017; Ramer et al., 2019). While surveys may be seen as the de-facto method to learn about recreationists' preferences, the ubiquitous use of the internet and social media has provided social and spatial scientists with a wealth of data from which inferences about preferences can be gained. Many social media platforms, such as the micro-blogging site Twitter and the photo-sharing site Flickr, make some or all of the data generated through posts on their platforms publicly available. Through the use of these data, researchers have been able to gain new insights into visitation numbers to parks and public lands, the geophysical characteristics (e.g., the presence of waterbodies, etc.) that are associated with visits, activities and preferences, and even the sentiment associated with different types of landscapes (Teles da Mota & Pickering, 2020; Wilkins, Wood, & Smith, 2021). While the rapidly growing body of research using social media is impressive, little work has been done to compare the content of images on social media relative to preferences elucidated via more traditional methods. Surveys are one way to measure visitors' *stated* preferences, while social media data are collected passively and are more likely to represent *revealed* preferences (Adamowicz, Louviere, & Williams, 1994). Although visitors might not only take photos of features they prefer (i.e., visitors may take photos of features they find interesting or unique, but not necessarily prefer), social media images could still be an indicator of visitor preferences (e.g., Hausmann et al., 2018; Väisänen, Heikinheimo, Hiippala, & Toivonen, 2021).

The purpose of this research is to compare the landscape features identified through a computer vision algorithm used to analyze social media photographs with preferences derived through a traditional on-site intercept survey. Our goal is to determine if a computer vision algorithm can generate valid inference about recreationists' landscape preferences. We address this goal by answering two research questions:

1. How do park visitors' social media photographs of landscape features compare to those that visitors say positively impacted their park experiences?
2. Do landscape preferences vary between visitors who post photographs of their park visit on social media and those who do not?

Our first question aims to critically assess challenges associated with using social media alone to generate inferences about landscape preferences. The second question tests the representativeness of social media users' preferences relative to non-users. This is a commonly cited limitation of social media (Wilkins et al., 2021), however no known studies have looked at potential differences in those who share photographs from experiences within parks and public lands and those who do not.

1.1. Literature review

Over the past decade, social media have proven to be a wellspring of new data capable of characterizing the use of parks and other public lands (e.g., Fisher et al., 2018; Ghermandi & Sinclair, 2019; Wood et al., 2020). Social media generally refers to online content that is user-generated and also hosted by a service (e.g., Twitter, Flickr, etc.) that facilitates connections between individuals or groups (Obar & Wildman, 2015). Social media can include text, photographs, and metadata such as the time a post was made or the location a photograph was taken. These pieces of information can be useful to social and spatial scientists interested in developing a better understanding of how parks and public lands are being used. To date, the majority of research using social media to understand recreation within parks and public lands has focused on evaluating the extent to which social media accurately represent the amount of visitation to a particular site (Teles da Mota & Pickering, 2020). Recent systematic reviews of this literature suggest social media does provide a relatively good (mean $r = 0.69$) indicator of the total volume of outdoor recreation occurring within parks and protected areas (Wilkins et al., 2021).

Aside from the use of social media as an indicator of the volume of use within parks and protected areas, it has also been used to characterize the spatial distribution of use and visitation across parks (Donahue et al., 2018; Hamstead et al., 2018; Heikinheimo et al., 2020; Kim, Kim, Lee, Lee, & Andrada, 2019; Li, Li, Li, & Long, 2020; Sinclair, Mayer, Woltering, & Ghermandi, 2020; Song, Richards, & Tan, 2020; Sonter, Watson, Wood, Ricketts, & Yang, 2016; Ullah et al., 2020; Zhang & Zhou, 2018) and within sub-regions of individual parks (Heikinheimo et al., 2017). These data have also been used to identify hot spots of use within individual parks or across broad geographic regions (Walden-Schreiner, Leung, & Tateosian, 2018; Walden-Schreiner, Rossi, Barros, Pickering, & Leung, 2018; Zhang, van Berkel, Howe, Miller, & Smith, 2021).

In addition to addressing questions of "how many" and "where" visitors are going, social media have been used to characterize the preferences of outdoor recreationists (Wilkins et al., 2021). This body of work has largely used social media to quantify recreationists' preferences for cultural ecosystem services. Distinct types of cultural ecosystem services in parks (e.g., recreation, aesthetic, scientific/educational, etc.) have most commonly been assessed through manual coding and classification of shared text or photographs (Clemente et al., 2019; Johnson, Campbell, Svendsen, & McMillen, 2019; Muñoz, Hausner, Runge, Brown, & Daigle, 2020; Retka et al., 2019; Van Berkel et al., 2018; Vaz et al., 2019; Vieira, Bragagnolo, Correia, Malhado, & Ladle, 2018). One study compared manual content analysis of social media photos in a national park to stated preferences for biodiversity from a visitor survey and found preferences for large-bodied mammals were over-represented in social media images, but small bodied mammals, plants, and reptiles were under-represented in images compared to survey-derived preferences (Hausmann et al., 2018). The time required to manually code and classify textual or photographic content curtails many of the advantages of using large, crowdsourced datasets. Additionally, manual coding introduces the possibility of researcher-introduced bias into the analysis (Araujo, Lock, & van de Velde, 2020).

A small body of research has attempted to overcome the bottleneck of manually coding social media images through the use of computer vision algorithms. Gosal, Geijzendorffer, Václavík, Poulin, and Ziv (2019) classified photographs uploaded to the Flickr social media platform with the Google Cloud Vision algorithm. Their analysis used over 20,000 shared photographs taken within the Camargue region in Southern France. The Google Cloud Vision algorithm returns a set of descriptive terms corresponding to the content of individual photographs. Informed by extensive training datasets, the algorithm is able to detect faces, objects, landmarks, and other content within the images (Google, 2021). Gosal et al. (2019) used the terms returned from the algorithm as inputs in latent semantic analysis to estimate meaning similarities between the terms; meaning similarities were subsequently used to identify discrete clusters of recreationists within the region. Relatedly, Runge, Hausner, Daigle, and Monz (2020) used Google Cloud Vision to process and classify over 800,000 Flickr photographs taken in the Arctic to characterize how individuals were interacting with nature in the region. The authors manually classified the descriptive terms returned from Google Cloud Vision into one of two broad cultural ecosystem services, abiotic nature and biotic nature. Those terms classified as biotic nature were subsequently classified into sub-categories (e.g., wildlife, bird, plant, etc.). The authors used the classification of the algorithm-defined terms to describe "how and where people interact with nature." Another recent study used Google Cloud Vision to understand and map aesthetic value of the landscape in a national park in northern England (Gosal & Ziv, 2020). In addition to research that has used computer vision algorithms to classify recreationists and the demand for cultural ecosystem services, recent work has used these algorithms to classify protected areas (Ghermandi, Sinclair, Fichtman, & Gish, 2020) and understand how the content of images differs between national and international visitors (Väisänen et al., 2021). Similar work has been done in other geographic areas less-known for their quality of outdoor recreation amenities or appeal to tourists. For example, Flickr photographs and computer vision algorithms such as Google Cloud Vision have been used to classify images in countries (Taecharungroj & Mathayomchan, 2020a), large groups of cities (Taecharungroj & Mathayomchan, 2020b), and individual cities (Richards & Tunçer, 2018).

The literature that has paired social media with computer vision algorithms in an attempt to understand individuals' preferences, activities, or values is admittedly sparse. The literature that does exist, as may have been gleaned from the review above, has predominately focused on geographic regions so large that ground-truthing or validating preferences inferred from photographic content has been overlooked. Moving too quickly to analyze big data with advanced, and sometimes poorly understood, algorithms may lead to an inaccurate or subjective understanding of human preferences or activities. More focused analysis on smaller geographic regions where comparisons can be drawn against 'tried and true' methods of understanding preferences are warranted to better understand the potential uses and limitations of social media image content and computer vision algorithms. In this investigation, we critically evaluate how photo content derived from social media and a computer vision algorithm compare to preferences derived from visitor surveys in an urban-proximate park system.

2. Methods

2.1. Study area

Boulder is a city in Colorado, USA that is about 40 km Northwest of Denver. This city has a department called Boulder Open Space and Mountain Parks (OSMP) that manages over 18,000 ha of parks and protected areas within and around the city. These lands contain over 250 km of developed trails and receive over six million visits a year (Leslie, 2018). Boulder is situated directly to the east of a mountain range, so the OSMP lands include both mountainous areas and flatlands with grass and prairies. Boulder OSMP managers have identified six distinctive 'landscape character areas' within their jurisdiction (Fig. 1). These areas include: (1) plains; (2) grasslands; (3) foothills; (4) peaks and unique topography; (5) remote lands; and (6) water. For the purposes of this study, we aggregated the plains landscape character area and the grasslands landscape character areas into a plains and grasslands landscape character area. We also aggregated the peaks and unique topography landscape character area with the remote lands landscape character area. These aggregations were made because of similarities in

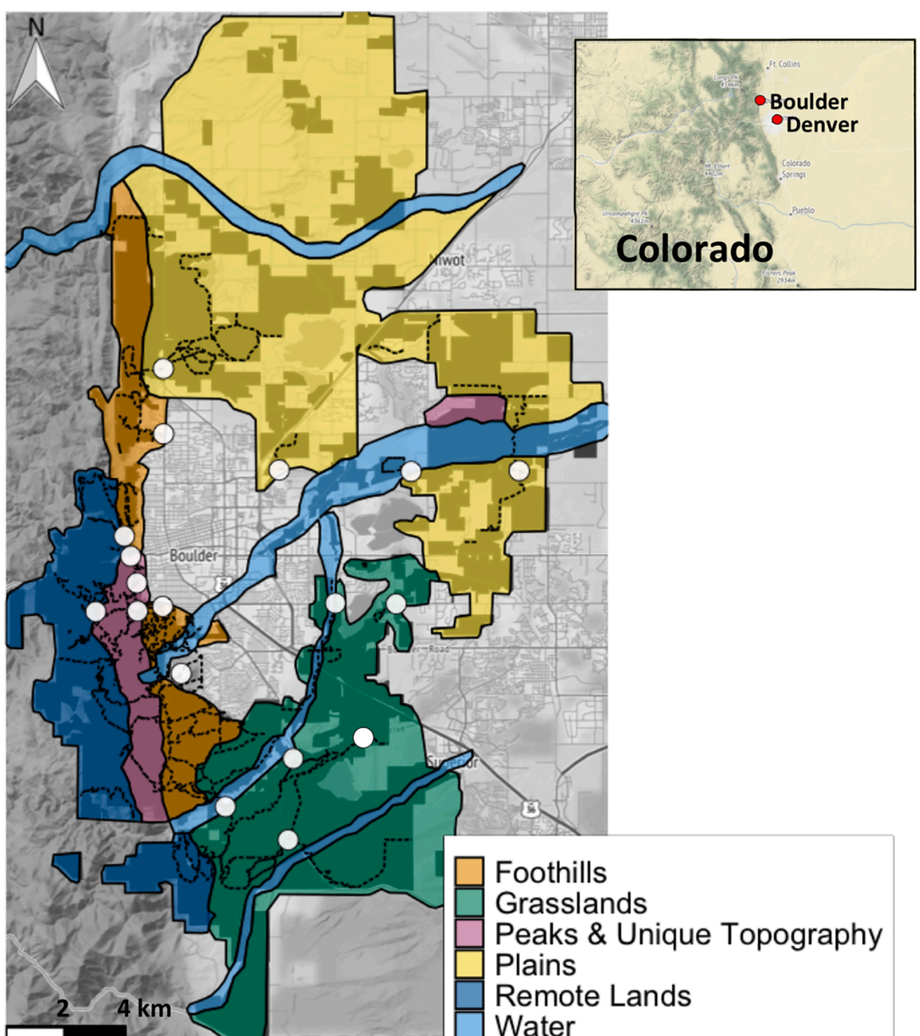

Fig. 1. Boulder Open Space and Mountain Parks lands, with the six different landscape character areas highlighted. White dots represent survey locations.

the aesthetic characteristics and landscape features within the combined areas. For example, both the plains and grasslands landscape character areas consist of flatlands without much woody vegetation (Dorning, van Berkel, Beck, Wilkins, Zhang, & Smith, 2019).

2.2. Data collection

2.2.1. Flickr data

We downloaded all Flickr images within Boulder OSMP boundaries directly through the platform's Application Programming Interface (API) using Python (Flickr, n.d.). These data represent images uploaded from 2004 to 2018 and were downloaded in April 2018. We ran all photographs through the Google Cloud Vision algorithm using the API (Google, 2021). For each photograph, the algorithm returns labels corresponding to the content of the image. The algorithm also provides a confidence score for each label. Google Cloud Vision only includes labels that have 50% confidence or higher. We used up to ten labels with the highest confidence scores to represent each photo, limiting our analysis to labels that had a confidence of 50% or higher. We downloaded the labels using the Google Cloud Vision API through Python.

2.2.2. Survey data

To determine recreationists' *stated* preferences for different landscape features, we collected data through an on-site survey. We provided a list of features frequently pictured in social media photographs on OSMP land and each respondent was asked to indicate how those features affected their experiences. Respondents were given five potential response options: *major negative impact*, *slight negative impact*, *neutral/did not see*, *slight positive impact*, or *major positive impact*. The ten features we asked visitors to rate were: (1) unique rock formations; (2) forested areas; (3) open plains and grasslands; (4) water; (5) old or historic buildings; (6) infrastructure; (7) development; (8) other people; (9) plants and other vegetation; and (10) agricultural land. We also asked about visitors' demographic characteristics, if they were planning to take photographs during their visit, if they would share them on social media, and if so, what platform. The full survey instrument is available as a supplementary file (Appendix A).

On-site questionnaires were distributed at 18 OSMP trailheads in May and June 2018. Survey locations were selected using a stratified sampling approach based on the six OSMP landscape character areas. We identified survey locations for targeted sampling using a spatial cluster analysis of geotagged Flickr posts and input from OSMP staff to refine sampling sites based on accessibility and use. Our cluster analysis approach identified the most prominent photographic clusters or scenic locations within each landscape character area. For each cluster, we identified the most popular trailhead providing access to the trails included within the cluster. See the supplementary material (Appendix B) for maps that display the spatial cluster analyses.

The 18 sampling locations were then randomly assigned to sampling days and times. We ensured each landscape character area was sampled at least twice on a weekday and at least once on a weekend (for a total of three sampling days each in the foothills, grasslands, plains, and water landscape character areas, and a total of four days each in the peaks and unique topography and remote lands areas). The sampling times were either in the morning (8 am to 2 pm) or afternoon (2 pm to 8 pm), and we only surveyed people over the age of 18. The adult in each group with a birthday closest to the day of the survey was selected to participate. A list of specific survey sites is available as a supplementary file (Appendix C), and the full database with survey results is publicly available (Wilkins & Smith, 2021).

2.3. Data analysis

We first acquired the frequency of each label returned from Google Cloud Vision across all Flickr photographs within Boulder OSMP. For any label that appeared 10 times or more, two authors independently coded these into the categories of landscape features asked about on the visitor survey. For example, labels such as "plant," "flora," "flower," and "flowering plant" were all coded as "plants and other vegetation." The two authors had a 94% agreement (491/523) in coding Google Vision labels into survey categories, and a third author served as a tiebreaker for the 6% of disagreements. For a full list of how labels were coded, see the supplementary material (Appendix D).

Each photograph was then categorized as representing one or more of the landscape features asked about in the survey based on that photograph's Google Vision labels. Each photograph could have multiple features. To determine the validity of using labels from Google Vision to automate the "viewing" of photographs for landscape features, one author manually viewed 5% of all Flickr images in this study. We categorized the Flickr images into the same ten landscape features asked about in the survey to determine the percent agreement between manually viewing photographs and using Google Vision labels to understand landscape features.

For the ten different landscape features, we then compared the percentages of people who said each landscape feature positively impacted their experience (via the on-site survey) to the percent of Flickr photographs in that landscape character area which contained that feature (via classified Google Vision labels). We ran Chi-Square tests of homogeneity to determine if there were differences between survey responses and Google Vision labels derived from the Flickr images. Chi-Square tests of homogeneity are used to determine whether two (or more) independent samples differ in their distributions on a single variable of interest (Franke, Ho, & Christie, 2012). In this case, we are looking at the presence or absence of each feature (e.g., forested areas) in Flickr photos compared to the percentage of surveyed visitors who said that same feature had a positive impact on their experience. We also used the Phi coefficient to understand the effect size, if differences existed. For this analysis, we used all Flickr photos in Boulder OSMP, compared to survey data (only collected in May and June). The photo content of May and June photos compared to all other months is shown in the supplementary material (Appendix E). Although there were some differences (e.g., more photos of plants and open plains and grasslands during May and June), all differences were very small (effect size < 0.1).

We then split visitors into those who were planning to share a photograph on social media from their current visit to OSMP, and those who were not planning to take or share photographs on social media. We ran Chi-Square tests of homogeneity between these two groups to determine if there were any differences in the impact of different landscape features on experiences depending on whether or not people were planning to take and share photographs on social media.

All analyses were conducted in *R* version 4.0.3 (R Core Team, 2020) and the code is publicly available (Wilkins & Smith, 2021).

3. Results

3.1. Sociodemographic characteristics of survey respondents

The overall response rate for the survey was 84.3%, with 537 respondents, and 100 refusals. Some questionnaires were not complete or usable ($n = 17$), so the final count of 520 represents an effective response rate of 81.6%. Of the questionnaires, 143 were from the plains and grasslands landscape character areas, 138 were from the foothills, 171 were from the peaks and unique topography and remote lands, and 68 were from water areas. Demographics of the survey sample are described in Table 1.

3.2. Landscape features identified by computer vision algorithm

In total, there were 9394 unique Flickr photographs from 885 users within Boulder OSMP lands. The distribution by landscape character area group is shown in Table 2. Across all landscape character areas, "mode of transport" was the most common label returned by the Google

Table 1
Demographics of the survey sample.

Characteristic and Responses	*n*	%
Lives in Boulder		
Yes	238	46.6
No	273	53.4
Age		
18–29	111	22.0
30–44	157	31.2
45–65	185	36.7
65+	51	10.1
Education		
Less than a Bachelor's	59	11.7
Bachelor's	222	43.9
Master's	141	27.9
Professional	27	5.3
Doctoral	57	11.3
Race/ethnicity		
White/Caucasian	452	91.5
Asian	31	6.3
Hispanic or Latino	24	4.9
African American	8	1.6
Other	4	0.8
Gender		
Male	259	51.2
Female	247	48.8
Other	0	0.0

Table 2
Flickr photograph characteristics by landscape character area.

Landscape character area	Flickr photographs	Unique users	Ten most frequent labels from Google Vision
Plains and grasslands	1,754	234	mode of transport (1,091) sky (9 2 1) ecosystem (9 0 7) land lot (4 5 5) prairie (4 2 5) grass (4 1 0) phenomenon (3 7 4) geological phenomenon (3 7 0) field (3 6 6) tree (3 6 0)
Foothills	2,543	444	mode of transport (1,629) ecosystem (1,327) sky (1,202) tree (9 4 9) geological phenomenon (8 2 6) land lot (6 4 3) phenomenon (6 0 3) grass (5 7 4) panorama (5 0 7) photography (4 9 8)
Peaks & unique topography and remote lands	4,026	463	mode of transport (2,972) geological phenomenon (2,152) tree (1,867) sky (1,846) ecosystem (1,790) phenomenon (1,436) panorama (7 9 0) photography (7 5 2) recreation (7 4 6) plant (6 2 6)
Water	1,071	137	mode of transport (5 6 7) tree (5 0 2) ecosystem (4 6 2) sky (3 5 3) grass (2 9 3) phenomenon (2 8 6) plant (2 7 1) geological phenomenon (2 6 3) land lot (2 3 4) water (2 2 8)

Vision algorithm (Table 2). Most of the top ten labels were the same across all landscape character areas. However, there were labels unique to some areas, such as "prairie" and "field" being in the top ten labels in the plains and grasslands areas, and "water" being in the top ten labels for the "water" area. Some labels, such as mode of transport, sky, and ecosystem, were not coded to represent any landscape feature (see Appendix D for how we coded Google Vision labels).

Overall, there was moderate agreement between manually coding a sample of the photographs, compared to how Google Vision labeled the photographs; the mean level of agreement across the ten landscape features was 78.6% (Table 3). Some categories had high agreement between manual codes and Google Vision labels, such as water (85.8%), infrastructure (84.9%), and agricultural land (92.9%). The largest discrepancies were in forested areas (63.2%) and plants/other vegetation (58.8%), where many more photographs were coded into these categories when manually viewing photographs compared to using Google Vision. This discrepancy indicates that forests and vegetation would be underrepresented when using Google Vision to classify landscape features within photographs.

We noticed that Google Vision had more trouble recognizing water that was muddy or brown, forests and features that were covered in snow, and development that was represented by lights (i.e., night scenes with buildings lit up in the distance). Google Vision seemed less likely to recognize there were people or forests in the image when individuals were more distant. Additionally, Google Vision could not accurately distinguish between a single tree and a forested area. We categorized any photo with the Google Vision label "tree" as a forested area, although when manually reviewing photos, sometimes this was not a forested area (e.g., possibly a plains area with other vegetation).

3.3. Data and method-dependent differences in preferred landscape features

There were significant differences between the percent of survey respondents who said each of the ten landscape features positively impacted their experience and the percent of Flickr photographs that depicted these same features (Table 4). Similar differences were found for nearly all of the landscape character areas within Boulder OSMP lands, where we observed significant differences in preferences comparing the ten surveyed landscape features, except for development features in the water landscape character area. Overall, visitors were more likely to say that landscape features had a positive impact on their experiences on surveys, but the features were not as likely to be depicted in an equivalent percentage of Flickr photographs. However, there is one exception to this, in that few visitors said development features had a positive impact on their experiences, while development features showed up more often in Flickr photographs. Fig. 2 displays the differences between the percentages of survey respondents who said each feature had a positive impact on their experience, compared to the percentage of Flickr photographs depicting each feature.

Effect sizes (Phi) varied, but most indicate medium to large differences in the percentages of survey respondents who said each feature had a positive impact on their experience, compared to the percentage of Flickr photographs. The largest differences were from old or historic buildings/structures, water, and plants and other vegetation being depicted less in Flickr photos compared to visitors' surveyed preferences.

Table 3
Agreement between landscape features identified by Google Vision (GV) relative to features identified manually for a random sample of Flickr photographs taken within the study area ($n = 437$).

	Agreement (%)	GV said no, but human said yes (%)	GV said yes, but human said no (%)	Difference in % of photos classified
Unique rock formations (stone slab, outcrops)	71.6	16.0	12.4	3.6 underrepresented by GV
Forested areas	63.2	30.0	6.9	23.1 underrepresented by GV
Open plains and grasslands	74.6	13.7	11.7	2.0 underrepresented by GV
Water (wetlands, lakes, streams)	85.8	4.4	9.8	5.4 overrepresented by GV
Old or historic buildings/ structures	98.2	0.5	1.4	0.9 overrepresented by GV
Infrastructure (fences, power lines, water tanks)	84.9	14.2	0.9	13.3 underrepresented by GV
Development (residential, industrial, commercial)	74.8	11.9	13.3	1.4 overrepresented by GV
Other people	81.0	15.3	3.7	11.6 underrepresented by GV
Plants and other vegetation	58.8	39.8	1.4	38.4 underrepresented by GV
Agricultural land	92.9	2.1	5.0	2.9 underrepresented by GV

3.4. Differences in landscape preferences between social media users and other visitors

The majority of visitors (61%; 318/519) said they were going to take photographs during their visit on the day they were intercepted. Of those, 75% (236/316) said they were planning to share the photographs on social media. The most popular platforms to share photographs from their visit were Instagram ($n = 155$) and Facebook ($n = 140$), with fewer visitors sharing photographs on Twitter ($n = 15$), Flickr ($n = 3$), or other platforms ($n = 27$). Visitors who were not from Boulder were more likely to share a photograph from their visit on social media (63% sharing a photograph) compared to those who were local residents (35% sharing a photograph).

There were no substantial differences in landscape preferences between those who were planning to upload a photograph from their visit to social media and those who were not going to post on social media (Table 5). There were statistically significant differences in preferences for unique rock formations, forested areas, development, and other people; however, the effect sizes (Phi) of these differences were all small. The visitors who were planning to upload photographs to social media were slightly more likely to indicate these features had a positive impact on their experiences.

4. Discussion

4.1. The use of computer vision to quantify landscape preferences

Surveys have been widely used for decades to understand visitor preferences in parks and protected areas (Manning, 2011). However, more research is starting to use social media to understand visitors' preferences (Wilkins et al., 2021). When comparing landscape preferences derived from surveys to the content of social media photographs, we found that most features appeared in a fewer percentage of Flickr photographs compared to the percentage of surveyed visitors that said the features positively impacted their experience. This is not surprising, considering many different features had positive impacts on visitors' experiences on Boulder OSMP parklands, and it is difficult to capture all of the important features in a single photograph. Development features were the only features that appeared in a larger percentage of photographs compared to survey-derived preferences. This is likely because OSMP is an urban-proximate parks system, so features such as cars, roads, and buildings were in the background of many photographs.

The large discrepancies we found between the content of Flickr photographs and preferences elucidated through on-site surveys suggests that caution needs to be taken before using social media and computer vision to understand landscape preferences. Our work echoes the findings of recent research that has compared the content of social media images in protected areas to online surveys, finding notable differences (Moreno-Llorca et al., 2020), and research that has compared visitor activities from surveys to social media photographs, again finding differences (Heikinheimo et al., 2017). Method-dependent findings are not unheard of in previous research on landscape preferences (e.g., Komossa, Wartmann, Kienast, & Verburg, 2020). In instances where different data sources and methods give incongruent results, the discrepancies may be attributable to different types of data and analytical methods capturing different user groups, different types of landscape features, or real differences between stated and revealed preferences (Adamowicz et al., 1994; Komossa et al., 2020; Wartmann, Acheson, & Purves, 2018). For instance, we surveyed visitors at trailheads, indicating most visitors in our survey sample were going for a walk, hike, or run. Some visitors in the Flickr sample may have just stopped to take a photo along a road or in a parking lot, representing a different type of visitor. In the case of social media broadly, some activities or landscape features may be important to visitors, but hard to capture in a photograph. Additionally, some photographs may have represented features visitors found unique or noteworthy, but not necessarily a preference. On the survey, visitors were only asked to rate if a feature has a positive or negative impact if they saw that feature on their current visit (i.e., they hypothetically could have taken a photo of that feature). Regardless, some visitors may have indicated that a feature would have a positive impact on their experience, even if they did not see it during that visit. Consequently, we suggest a diversity of data sources and analytical methods should be used in complementary and comparative ways. Our analysis here suggests both surveys and social media photographs can provide important insights about landscape preferences, but neither in isolation is perfect.

We did, however, find that computer vision was useful to automate the content analysis of a large number of photographs, and that using Google Cloud Vision to automate photograph classification is fairly consistent with manually coding images for some features (e.g., water, infrastructure), but not as consistent for other features (e.g., plants, forests). When differences in classification did exist, the automated classification underrepresented landscape features compared to manual content analysis. Runge et al.'s analysis of Flickr photographs taken within the Arctic also found that Google Vision underrepresented biotic nature features compared to manual coding, and that computer vision was not great at identifying specific types of wildlife (2020). Even if studies use manual coding, there is still bias and subjectivity into how images are coded, and different researchers may not code the landscape features depicted in photos in the same way. Manual coding may be more appropriate for some features, while automated coding may be less biased for other features. The ability of computer vision algorithms to detect certain features depends on what images they are trained on, and these algorithms are likely to improve in the future.

Table 4
Comparison of survey respondents who said each landscape feature had a positive impact on their experience, to the percent of Flickr photographs that depicted each feature. Landscape features in Flickr photographs were identified from Google Vision; therefore, some categories might be underrepresented compared to manually classifying Flickr photographs.

	Overall			Plains and Grasslands			Peaks, unique topography, and remote lands			Foothills			Water		
	% positive impact	% of photo-graphs	X^2 (Phi)	% positive impact	% of photo-graphs	X^2 (Phi)	% positive impact	% of photo-graphs	X^2 (Phi)	% positive impact	% of photo-graphs	X^2 (Phi)	% positive impact	% of photo-graphs	X^2 (Phi)
Unique rock formations (stone slab, outcrops)	86.9	43.0	366.9 (0.19)	62.4	22.6	97.7 (0.23)	97.1	58.1	103.4 (0.16)	92.6	40.1	144.3 (0.23)	95.5	26.6	138.7 (0.35)
Forested areas	93.4	40.0	555.8 (0.24)	83.2	21.8	238.0 (0.36)	98.2	46.8	172.6 (0.20)	94.0	38.5	161.7 (0.25)	100.0	47.6	69.2 (0.25)
Open plains and grasslands	92.4	26.9	985.5 (0.32)	97.1	40.5	166.8 (0.30)	88.2	15.0	591.5 (0.38)	94.9	32.8	215.6 (0.28)	88.1	35.1	74.9 (0.26)
Water (wetlands, lakes, streams)	84.6	13.0	1780.7 (0.42)	95.0	13.1	561.2 (0.54)	81.3	12.3	596.1 (0.38)	76.1	8.1	578.4 (0.47)	88.2	27.2	111.8 (0.31)
Old or historic buildings/structures	45.2	1.6	2426.0 (0.50)	41.1	0.3	654.6 (0.59)	41.1	2.5	606.2 (0.38)	51.9	1.5	840.6 (0.56)	49.3	0.7	439.2 (0.62)
Infrastructure (fences, power lines, water tanks)	12.0	2.9	116.7 (0.11)	14.6	5.1	20.0 (0.10)	11.7	2.5	46.5 (0.11)	8.3	2.0	21.9 (0.09)	14.9	3.1	24.3 (0.15)
Development (residential, industrial, commercial)	11.8	21.6	26.5 (0.05)	10.1	18.8	6.2 (0.06)	14.8	23.9	7.2 (0.04)	9.9	20.7	9.0 (0.06)	11.8	19.2	2.3 (n/s)
Other people	53.9	21.5	284.5 (0.17)	55.8	12.9	175.2 (0.30)	59.3	25.1	96.6 (0.15)	44.1	25.2	23.8 (0.09)	56.7	13.4	87.9 (0.28)
Plants and other vegetation	95.3	20.9	1431.4 (0.38)	95.0	16.1	467.6 (0.50)	94.0	20.9	469.2 (0.34)	97.8	20.4	419.0 (0.40)	94.0	29.9	116.2 (0.32)
Agricultural land	51.7	5.9	1315.6 (0.37)	79.7	13.1	374.5 (0.45)	37.0	1.4	734.8 (0.42)	32.8	8.0	91.4 (0.19)	68.7	5.7	293.4 (0.51)
Sample sizes (*n*)	497–511	9,394		124–140	1,754		162–171	4,026		131–136	2,543		66–68	1,071	

Note. All Chi-Square tests are significant at $\alpha < 0.05$ (except the differences for development features in the water landscape character) indicating significant differences between the proportion of survey respondents who said each landscape feature had a positive impact on their experience and the percent of Flickr photographs that depicted each feature.

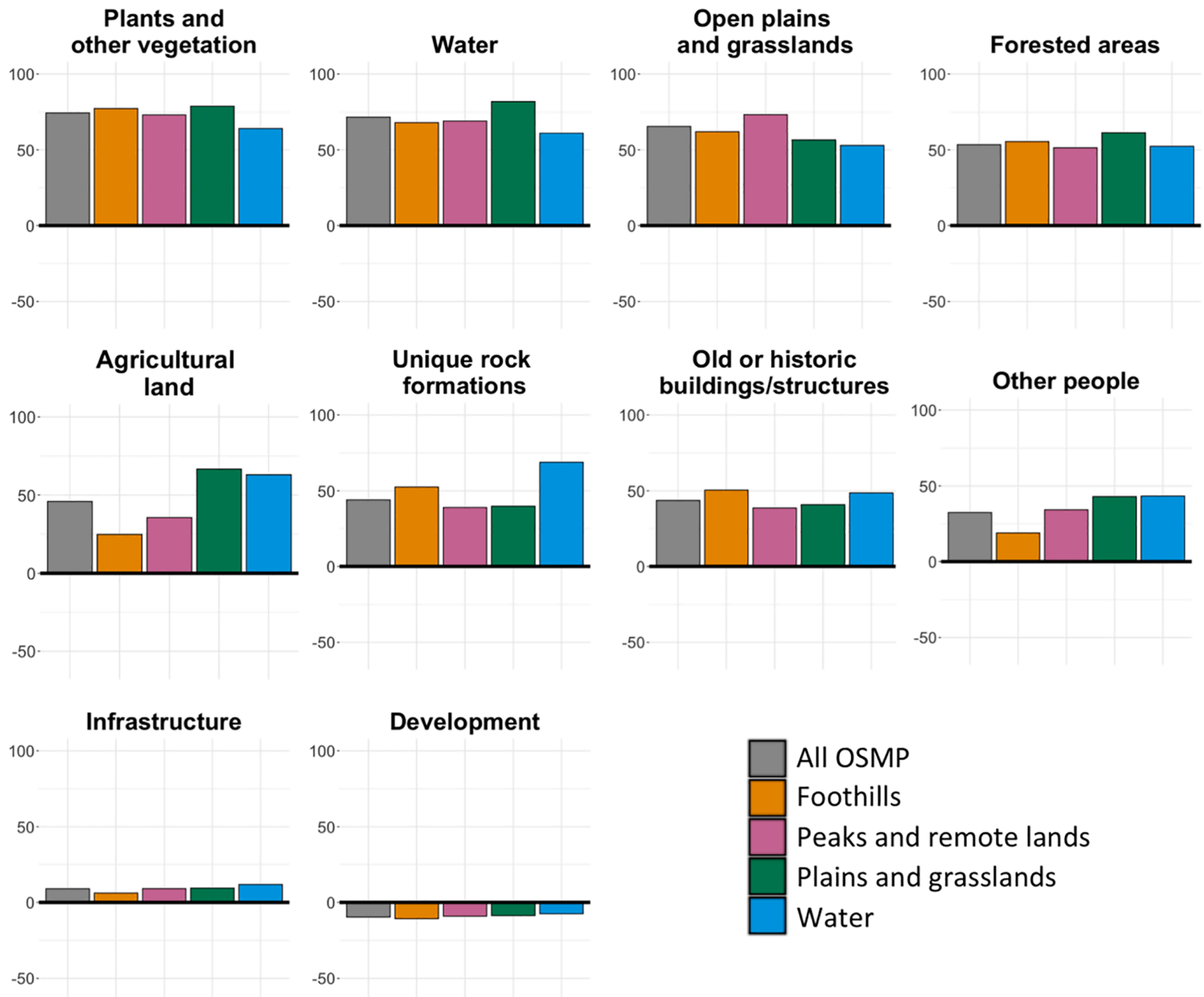

Fig. 2. Differences between the percentages of survey respondents who said each feature had a positive impact on their experience and the percentages of Flickr photographs that depicted each feature. Positive values represent more surveyed visitors saying the feature had a positive impact on their experience, while negative values represent a greater representation in Flickr photographs.

Given that landscape preference research may increasingly be turning to computer vision algorithms to understand individuals' preferences and their experiences within parks and public lands, more research is needed to better understand what landscape features different computer vision algorithms are good at identifying, and which features might be better identified manually. A recent study found that different computer vision algorithms produced different results, with some algorithms being better for certain features. For example, the authors found that Google Vision was preferable for identifying features like wildlife and vegetation, while Clarifai was better at distinguishing feelings and sentiment (Ghermandi, Depietri, & Sinclair, 2022). Future research may consider using a scaling-factor to account for underrepresentation of landscape features that are poorly captured using certain computer vision algorithms (e.g., we found the Google Vision algorithm poorly captured plants and other vegetation), or using different algorithms for different purposes. One recent study used different computer-vision methods and algorithms to serve different and specific purposes, for instance, classifying scenes or identifying objects (Väisänen et al., 2021). This approach may be preferable to using only one method and algorithm, but more research is needed to see how different algorithms compare in their ability to accurately assess landscape features (Ghermandi et al., 2022).

4.2. *The representativeness of social media users to other visitors to parks and public lands*

Although a commonly stated limitation of data derived from social media is that it may not be representative of all park visitors (Wilkins et al., 2021), we did not find any substantial differences between the landscape preferences of visitors who self-reported posting photographs of their visit to OSMP lands on social media and those who said they would not post photographs. Although we did find that visitors who said they would post images to social media were slightly more positive about seeing rock formations, forests, development, and other people, the effect size of the differences were all small. This finding suggests using social media data may be representative to answer certain types of questions, but more research is needed to determine if this holds true in other geographic locations and for other types of research questions (e.g., understanding sentiment or behavior).

Table 5
Differences in landscape feature preferences between those who plan to post photographs of their visit on social media, and those who do not.

	Will upload to social media: % positive impact	Will NOT upload to social media: % positive impact	X^2(Phi)	p-value
Unique rock formations (stone slab, outcrops)	91.7	82.8	8.7 (0.13)	0.003
Forested areas	96.5	90.8	6.7 (0.12)	0.010
Open plains and grasslands	91.9	92.8	0.1 (n/s)	0.712
Water (wetlands, lakes, streams)	84.8	84.5	0.0 (n/s)	0.924
Old or historic buildings/ structures	43.2	46.9	0.7 (n/s)	0.415
Infrastructure (fences, power lines, water tanks)	11.6	12.3	0.1 (n/s)	0.810
Development (residential, industrial, commercial)	15.2	9.1	4.4 (0.09)	0.037
Other people	60.7	48.6	7.5 (0.12)	0.006
Plants and other vegetation	95.2	95.3	0.0 (n/s)	0.978
Agricultural land	49.3	53.6	0.9 (n/s)	0.350
Sample sizes (*n*)	222–234	262–278		

4.3. Limitations

One limitation of this study is that we only analyzed photographs from the Flickr photo-sharing platform. Photograph content posted on other social media platforms might differ (Ghermandi et al., 2020; Hausmann et al., 2018). We also analyzed all Flickr photographs within Boulder OSMP, which indicates that the content of users who post many photographs was likely overrepresented. We only surveyed one visitor (an adult) in each group in Boulder OSMP, but Flickr could include photos taken by people under 18, or from multiple people within a group. Additionally, we only surveyed visitors in May and June, but looked at photograph content from all seasons, and there are some small differences in photo content taken during May and June compared to the rest of the year. Finally, additional research could help better understand how social media photograph content compares to surveys in other contexts and locations (e.g., parks that are not urban-proximate).

4.4. Implications for planning and management

Aside from the methodologically oriented findings of this work, the analysis does also provide some insights for Boulder OSMP managers. We found a large majority of visitors to Boulder OSMP reported that natural features such as unique rock formations, forested areas, open plains and grasslands, water, and plants had a positive impact on their experience. Seeing infrastructure and development was only positive for a small portion of visitors (12% and 15%, respectively). This suggests that visitors may seek out trails and places within OSMP that are perceived as more natural and have less infrastructure and development visible. These preferences could be considered if OSMP were to purchase additional parklands in the future or add new trails. Additionally, we found the majority of people said that seeing other people had a positive impact on their experience. Although increasing use is often seen as a negative influence on park visitors' experiences (Manning, Valliere, Minteer, Wang, & Jacobi, 2000), seeing some other people enhances the experience of many OSMP visitors. Until there is further research on using social media to infer visitors' preferences, we would recommend that land managers and planners continue to use visitor surveys to glean preferences rather than relying on social media data alone.

5. Conclusions

Social media has transformed how many individuals experience parks and public lands. It has also reshaped how those experiences are shared with friends, family, and (now) followers. The rapid ascent in the popularity of social media has led social and spatial scientists to rethink how they investigate individuals' preferences within parks and other public lands. The research using social media to understand the preferences of outdoor recreationists and tourists is growing at an outstanding pace (Wilkins et al., 2021). This brings up the question: can the photographs that individuals share on social media be used to infer preferences? And are the preferences and behaviors of those who use social media representative of all visitors to parks and public lands? These are important methodological questions that deserve substantial attention before the field embraces findings from new and novel data sources and analytical methods.

In this investigation, we found that while the Google Cloud Vision algorithm was relatively good at classifying landscape features (mean agreement of 78.6% relative to manual coding), there were consistent and significant differences in the features identified by the algorithm and those features that park users said improved their recreational experiences. At the least, caution should be taken in interpreting landscape preferences from computer vision algorithms and social media. This is not to say that computer vision algorithms and social media cannot be used as valuable tools to elucidate preferences, perceptions, and other aspects of the outdoor recreation experience. These algorithms can be useful at quantifying *broad* types of visitor experiences. For example, previous work has found success in using computer vision to characterize human-wildlife interactions (Runge et al., 2020). Trying to infer preferences for specific types of landscape features may be too fine grained of a research question for current computer vision algorithms. This may change in the future as these algorithms are further developed and refined. In short, our analysis highlights the need to be cautious when using computer vision algorithms and social media data to draw inferences about landscape preferences. Multiple data sources and analytical methods are warranted to provide a check on the promises of computer vision algorithms and social media and avoid the pitfalls that can accompany their use.

CRediT authorship contribution statement

Emily J. Wilkins: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – Original draft, Writing – Review & editing, Visualization. **Derek Van Berkel:** Conceptualization, Software, Writing – Review & editing. **Hongchao Zhang:** Investigation, Writing – Review & editing. **Monica A. Dorning:** Conceptualization, Writing – Review & editing. **Scott M. Beck:** Conceptualization, Writing – Review & editing. **Jordan W. Smith:** Conceptualization, Resources, Writing – Original draft, Writing – Review & editing, Supervision, Project administration.

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

Funding for this research was provided by Boulder Open Space and Mountain Parks and by the Institute of Outdoor Recreation and Tourism at Utah State University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2021.104315.

References

Adamowicz, W., Louviere, J., & Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management, 26*(3), 271–292. https://doi.org/10.1006/jeem.1994.1017

Araujo, T., Lock, I., & van de Velde, B. (2020). Automated visual content analysis (AVCA) in communication research: A protocol for large scale image classification with pre-trained computer vision models. *Communication Methods and Measures, 14*(4), 239–265. https://doi.org/10.1080/19312458.2020.1810648

Clemente, P., Calvache, M., Antunes, P., Santos, R., Cerdeira, J. O., & Martins, M. J. (2019). Combining social media photographs and species distribution models to map cultural ecosystem services: The case of a Natural Park in Portugal. *Ecological Indicators, 96*, 59–68. https://doi.org/10.1016/j.ecolind.2018.08.043

DeLucio, J., & Múgica, M. (1994). Landscape preferences and behaviour of visitors to Spanish national parks. *Landscape and Urban Planning, 29*(2–3), 145–160. https://doi.org/10.1016/0169-2046(94)90024-8

Donahue, M. L., Keeler, B. L., Wood, S. A., Fisher, D. M., Hamstead, Z. A., & McPhearson, T. (2018). Using social media to understand drivers of urban park visitation in the Twin Cities, MN. *Landscape and Urban Planning, 175*, 1–10. https://doi.org/10.1016/j.landurbplan.2018.02.006

Dorning, M. A., van Berkel, D., Beck, S. M., Wilkins, E. J., Zhang, H., & Smith, J. W. (2019). *Aesthetic characteristics of the Front Range: An analysis of viewsheds provided by Boulder OSMP lands*. Institute of Outdoor Recreation and Tourism, Utah State University. https://digitalcommons.usu.edu/extension_curall/1975/.

Fisher, D. M., Wood, S. A., White, E. M., Blahna, D. J., Lange, S., Weinberg, A., … Lia, E. (2018). Recreational use in dispersed public lands measured using social media data and on-site counts. *Journal of Environmental Management, 222*, 465–474. https://doi.org/10.1016/j.jenvman.2018.05.045

Flickr. (n.d.). *The Flickr developer guide: API*. https://www.flickr.com/services/developer/api/.

Franke, T. M., Ho, T., & Christie, C. A. (2012). The chi-square test: Often used and more often misinterpreted. *American Journal of Evaluation, 33*(3), 448–458. https://doi.org/10.1177/1098214011426594

Ghermandi, A., Depietri, Y., & Sinclair, M. (2022). In the AI of the beholder: A comparative analysis of computer vision-assisted characterizations of human-nature interactions in urban green spaces. *Landscape and Urban Planning, 217*, 104261. https://doi.org/10.1016/j.landurbplan.2021.104261

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

Ghermandi, A., Sinclair, M., Fichtman, E., & Gish, M. (2020). Novel insights on intensity and typology of direct human-nature interactions in protected areas through passive crowdsourcing. *Global Environmental Change, 65*, 102189. https://doi.org/10.1016/j.gloenvcha.2020.102189

Google. (2021). *Cloud Vision API*. https://cloud.google.com/vision.

Gosal, A. S., Geijzendorffer, I. R., Václavík, T., Poulin, B., & Ziv, G. (2019). Using social media, machine learning and natural language processing to map multiple recreational beneficiaries. *Ecosystem Services, 38*, 100958. https://doi.org/10.1016/j.ecoser.2019.100958

Gosal, A. S., & Ziv, G. (2020). Landscape aesthetics: Spatial modelling and mapping using social media images and machine learning. *Ecological Indicators, 117*, 106638. https://doi.org/10.1016/j.ecolind.2020.106638

Hamstead, Z. A., Fisher, D., Ilieva, R. T., Wood, S. A., McPhearson, T., & Kremer, P. (2018). Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers, Environment and Urban Systems, 72*, 38–50. https://doi.org/10.1016/j.compenvurbsys.2018.01.007

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., & Di Minin, E. (2018). Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. *Conservation Letters, 11*(1), e12343. https://doi.org/10.1111/conl.12343

Heikinheimo, V., Minin, E. D., Tenkanen, H., Hausmann, A., Erkkonen, J., & Toivonen, T. (2017). User-generated geographic information for visitor monitoring in a national park: A comparison of social media data and visitor survey. *ISPRS International Journal of Geo-Information, 6*(3), 85. https://doi.org/10.3390/ijgi6030085

Heikinheimo, V., Tenkanen, H., Bergroth, C., Järv, O., Hiippala, T., & Toivonen, T. (2020). Understanding the use of urban green spaces from user-generated geographic information. *Landscape and Urban Planning, 201*, 103845. https://doi.org/10.1016/j.landurbplan.2020.103845

Howley, P. (2011). Landscape aesthetics: Assessing the general publics' preferences towards rural landscapes. *Ecological Economics, 72*, 161–169. https://doi.org/10.1016/j.ecolecon.2011.09.026

Jiang, Y., & Yuan, T. (2017). Public perceptions and preferences for wildflower meadows in Beijing, China. *Urban Forestry & Urban Greening, 27*, 324–331. https://doi.org/10.1016/j.ufug.2017.07.004

Johnson, M. L., Campbell, L. K., Svendsen, E. S., & McMillen, H. L. (2019). Mapping urban park cultural ecosystem services: A comparison of Twitter and semi-structured interview methods. *Sustainability, 11*(21), 6137. https://doi.org/10.3390/su11216137

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology, 15*(3), 169–182. https://doi.org/10.1016/0272-4944(95)90001-2

Kim, Y., Kim, C., Lee, D. K., Lee, H., & Andrada, R. II. T. (2019). Quantifying nature-based tourism in protected areas in developing countries by using social big data. *Tourism Management, 72*, 249–256. https://doi.org/10.1016/j.tourman.2018.12.005

Komossa, F., Wartmann, F. M., Kienast, F., & Verburg, P. H. (2020). Comparing outdoor recreation preferences in peri-urban landscapes using different data gathering methods. *Landscape and Urban Planning, 199*, 103796. https://doi.org/10.1016/j.landurbplan.2020.103796

Leslie, C. (2018). *2016-2017 Visitation estimate report*. The City of Boulder, Open Space and Mountain Parks Department. https://www-static.bouldercolorado.gov/docs/2016-2017_VisitationEstimateReport_v4.0-1-201811141330.pdf?_ga=2.143164919.376274620.1612914381-431470235.1612914381.

Li, F., Li, F., Li, S., & Long, Y. (2020). Deciphering the recreational use of urban parks: Experiments using multi-source big data for all Chinese cities. *Science of The Total Environment, 701*, 134896. https://doi.org/10.1016/j.scitotenv.2019.134896

Manning, R. (2011). *Studies in outdoor recreation: Search and research for satisfaction* (3rd ed.). Oregon State University Press.

Manning, R., Valliere, W. A., Minteer, B., Wang, B., & Jacobi, C. (2000). Crowding in parks and outdoor recreation: A theoretical, empirical, and managerial analysis. *Journal of Park and Recreation Administration, 18*, 57–72.

Moreno-Llorca, R., Méndez, P. F., Ros-Candeira, A., Alcaraz-Segura, D., Santamaría, L., Ramos-Ridao, Á. F., … Vaz, A. S. (2020). Evaluating tourist profiles and nature-based experiences in Biosphere Reserves using Flickr: Matches and mismatches between online social surveys and photo content analysis. *Science of The Total Environment, 737*, 140067. https://doi.org/10.1016/j.scitotenv.2020.140067

Múgica, M., & De Lucio, J. V. (1996). The role of on-site experience on landscape preferences. A case study at Doñana National Park (Spain). *Journal of Environmental Management, 47*(3), 229–239.

Muñoz, L., Hausner, V. H., Runge, C., Brown, G., & Daigle, R. (2020). Using crowdsourced spatial data from Flickr vs. PPGIS for understanding nature's contribution to people in Southern Norway. *People and Nature, 2*(2), 437–449. https://doi.org/10.1002/pan3.10083

Obar, J. A., & Wildman, S. S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy, 39*(9), 745–750. https://doi.org/10.2139/ssrn.2647377

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ramer, H., Nelson, K. C., Spivak, M., Watkins, E., Wolfin, J., & Pulscher, M. (2019). Exploring park visitor perceptions of 'flowering bee lawns' in neighborhood parks in Minneapolis, MN, US. *Landscape and Urban Planning, 189*, 117–128. https://doi.org/10.1016/j.landurbplan.2019.04.015

Retka, J., Jepson, P., Ladle, R. J., Malhado, A. C. M., Vieira, F. A. S., Normande, I. C., … Correia, R. A. (2019). Assessing cultural ecosystem services of a large marine protected area through social media photographs. *Ocean & Coastal Management, 176*, 40–48. https://doi.org/10.1016/j.ocecoaman.2019.04.018

Richards, D. R., & Tunçer, B. (2018). Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem Services, 31*, 318–325. https://doi.org/10.1016/j.ecoser.2017.09.004

Rosenberger, R. S., Bergerson, T. R., & Kline, J. D. (2009). Macro-linkages between health and outdoor recreation: The role of parks and recreation providers. *Journal of Park and Recreation Administration. 27 (3): 8-20.*, *27*(3), 8-20.

Runge, C. A., Hausner, V. H., Daigle, R. M., & Monz, C. A. (2020). Pan-Arctic analysis of cultural ecosystem services using social media and automated content analysis. *Environmental Research Communications, 2*(7), Article 075001. https://doi.org/10.1088/2515-7620/ab9c33

Sinclair, M., Mayer, M., Woltering, M., & Ghermandi, A. (2020). 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

Song, X. P., Richards, D. R., & Tan, P. Y. (2020). Using social media user attributes to understand human–environment interactions at urban parks. *Scientific Reports, 10*(1), 808. https://doi.org/10.1038/s41598-020-57864-4

Sonter, L. J., Watson, K. B., Wood, S. A., Ricketts, T. H., & Yang, J. (2016). Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLOS ONE, 11*(9), e0162372.

Stein, T. V., & Lee, M. E. (1995). Managing recreation resources for positive outcomes: An application of benefits-based management. *Journal of Park and Recreation Administration, 13*(3), 52–70.

Taecharungroj, V., & Mathayomchan, B. (2021). Traveller-generated destination image: Analysing Flickr photos of 193 countries worldwide. *International Journal of Tourism Research, 23*(3), 417–441. https://doi.org/10.1002/jtr.2415

Taecharungroj, V., & Mathayomchan, B. (2020b). The big picture of cities: Analysing Flickr photos of 222 cities worldwide. *Cities, 102*, 102741. https://doi.org/10.1016/j.cities.2020.102741

Teles da 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*, Article 100295. https://doi.org/10.1016/j.jort.2020.100295

Ullah, H., Wan, W., Haidery, S. A., Khan, N. U., Ebrahimpour, Z., & Muzahid, A. A. M. (2020). Spatiotemporal patterns of visitors in urban green parks by mining social media big data based upon WHO reports. *IEEE Access, 8*, 39197–39211. https://doi.org/10.1109/ACCESS.2020.2973177

Väisänen, T., Heikinheimo, V., Hiippala, T., & Toivonen, T. (2021). Exploring human–nature interactions in national parks with social media photographs and computer vision. *Conservation Biology, 35*(2), 424–436. https://doi.org/10.1111/cobi.13704

Van Berkel, D. B., Tabrizian, P., Dorning, M. A., Smart, L., Newcomb, D., Mehaffey, M., … Meentemeyer, R. K. (2018). Quantifying the visual-sensory landscape qualities that contribute to cultural ecosystem services using social media and LiDAR. *Ecosystem Services, 31*, 326–335. https://doi.org/10.1016/j.ecoser.2018.03.022

Vaz, A. S., Gonçalves, João. F., Pereira, P., Santarém, F., Vicente, J. R., & Honrado, João. P. (2019). Earth observation and social media: Evaluating the spatiotemporal contribution of non-native trees to cultural ecosystem services. *Remote Sensing of Environment, 230*, 111193. https://doi.org/10.1016/j.rse.2019.05.012

Velarde, M. D., Fry, G., & Tveit, M. (2007). Health effects of viewing landscapes – Landscape types in environmental psychology. *Urban Forestry & Urban Greening, 6* (4), 199–212. https://doi.org/10.1016/j.ufug.2007.07.001

Vieira, F. A. S., Bragagnolo, C., Correia, R. A., Malhado, A. C. M., & Ladle, R. J. (2018). A salience index for integrating multiple user perspectives in cultural ecosystem service assessments. *Ecosystem Services, 32*, 182–192. https://doi.org/10.1016/j.ecoser.2018.07.009

Walden-Schreiner, C., Leung, Y.-F., & Tateosian, L. (2018). Digital footprints: Incorporating crowdsourced geographic information for protected area management. *Applied Geography, 90*, 44–54. https://doi.org/10.1016/j.apgeog.2017.11.004

Walden-Schreiner, C., Rossi, S. D., Barros, A., Pickering, C., & Leung, Y.-F. (2018). Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas. *Ambio, 47*(7), 781–793. https://doi.org/10.1007/s13280-018-1020-4

Wang, R., Zhao, J., Meitner, M. J., Hu, Y., & Xu, X. (2019). Characteristics of urban green spaces in relation to aesthetic preference and stress recovery. *Urban Forestry & Urban Greening, 41*, 6–13. https://doi.org/10.1016/j.ufug.2019.03.005

Wartmann, F. M., Acheson, E., & Purves, R. S. (2018). Describing and comparing landscapes using tags, texts, and free lists: An interdisciplinary approach. *International Journal of Geographical Information Science, 32*(8), 1572–1592. https://doi.org/10.1080/13658816.2018.1445257

Wilkins, E. J., & Smith, J. W. (2021). Promises and pitfalls of using computer vision to make inferences about landscape preferences: Evidence from an urban-proximate park system (data and code). Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-10-31. doi:10.3886/E139681V1.

Wilkins, E. J., Wood, S. A., & Smith, J. W. (2021). Uses and limitations of social media to inform visitor use management in parks and protected areas: A systematic review. *Environmental Management, 67*(1), 120–132. https://doi.org/10.1007/s00267-020-01373-7

Wood, S. A., Winder, S. G., Lia, E. H., White, E. M., Crowley, C. S., & Milnor, A. A. (2020). Next-generation visitation models using social media to estimate recreation on public lands. *Scientific Reports, 10*(1), 1–12. https://doi.org/10.1038/s41598-020-70829-x

Zhang, H., van Berkel, D., Howe, P. D., Miller, Z. D., & Smith, J. W. (2021). Using social media to measure and map visitation to public lands in Utah. *Applied Geography, 128*, 102389. https://doi.org/10.1016/j.apgeog.2021.102389

Zhang, S., & Zhou, W. (2018). Recreational visits to urban parks and factors affecting park visits: Evidence from geotagged social media data. *Landscape and Urban Planning, 180*, 27–35. https://doi.org/10.1016/j.landurbplan.2018.08.004