Preprints.org

Article

Identifying Urban Park Events through Computer Vision-Assisted Categorization of Publicly-Available Imagery

Yizhou Tan , Wenjing Li , Da Chen , Waishan Qiu [*]

Article

Identifying Urban Park Events through Computer Vision-Assisted Categorization of Publicly-Available Imagery

Yizhou Tan [1], Wenjing Li [2], Da Chen [3] and Waishan Qiu [4,*

[1] Department of Architecture, Rhode Island School of Design, Rhode Island/USA; ytan02@risd.edu
[2] Center for Spatial Information Science, The University of Tokyo, Tokyo/Japan; liwenjing@csis.u-tokyo.ac.jp
[3] Department of Computer Science, The University of Bath, Bath/UK; da.chen@bath.edu
[4] Department of City and Regional Planning, Cornell University, New York/USA; wq43@cornell.edu
* Correspondence: wq43@cornell.edu, qiuwaishan@126.com; Tel.: +86 13916556218

Abstract: Understanding park events and their categorization offer pivotal insights into urban parks and their integral roles in cities. This study utilized images and event category data from the New York City Parks Events Listing database to train a Convolutional Neural Network (CNN) for image-based park event categorization. Different CNN models were tuned to complete this multi-label classification task, their performances compared. Preliminary results underscore the efficacy of deep learning in automating the event classification process, revealing the multifaceted activities within urban green spaces. The CNN showcased proficiency in discerning various event nuances, emphasizing the diverse recreational and cultural offerings of urban parks. Such categorization has potential applications in urban planning, aiding decision-making processes related to resource distribution, event coordination, and infrastructure enhancements tailored to specific park activities.

Keywords: urban green spaces; human activities; Convolutional Neural Networks

1. Introduction

1.1. Background

Urban parks play a vital role in cities and their significance has continuously evolved in the lives of city dwellers. The benefits of urban parks include environmental benefits such as biodiversity and local cooling, economic benefits such as energy savings and property value, and social and psychological benefits such as physical activity and reduced obesity [1][2]. One of the important topics of park-related research is humans' events and programs in parks. Many studies have shown how park events could become a deciding force in shifting the park's own functionality [3][4][5][6]. In a report investigating London's urban parks, Smith and Vodicka [3] summarized from accounts of Friends groups that events are seen as a promotion of the park's inclusivity that brings more people into the park, contributing to community cohesion. A similar study by Neal et al. [6] on parks also credits urban park events as an opportunity of inclusivity, as organized events present a more ethnically diverse population than regular park users. Citroni and Karrholm [7] mentioned the relationship of events to civility, and how events facilitate the visibility of everyday life and forged a pattern of urban civility. Studying events in urban parks provides an insight for us to understand how these parks could actively contribute to a city and its community and help us reach a more sustainable city with high quality of life.

There is a significant gap between existing works and efficient event analysis of the parks. First, most of the past studies about park event analysis have been focusing on the intensity of park use, demographics of park users, the periods of time parks are used, and the level of physical activities. However, few studies have been focusing on the categorization of park events and programs, Secondly, from the aspect of data source, a majority of current studies analyzing the categories of human activities and planned events in parks have relied on mass questionnaires and interviews

[8][9][10][11], which is time consuming and site restrictive. Recent technological methods introduce big data into detailed park use analysis, such as GPS data and public participation geographic information systems data. However, GPS-based mobile phone tracking is not informative to the categorization of events and recreational park use [12], and public participation geographic information systems (PPGIS) cannot guarantee data sufficiency [13]. Social media data and other publicly available online imagery are a good source of information regarding recreational use of parks. Thirdly, from the aspect of methodology the methods of existing studies are either inefficient or not specifically targeted towards park events. Recent studies that utilize publicly available online imagery still involve tedious manual classifications [12]. The current research status calls for an updated methodology of a more accessible and cost-effective urban park events category analysis.

With the New York City Parks Events Listing [14] data which is a set of publicly available, tagged image data, this study proposes an algorithm featuring deep learning methods to more efficiently identify events and programming in urban parks by analyzing publicly available images of these parks, and performing classification based on park events. This is for the purpose of helping urban researchers and planners to better understand the impacts of park events in the community, and further incorporate them into the decision-making process.

1.2. Related Works

Although a significant number of studies have been conducted to determine the use of urban parks, the majority of these studies focused quantitatively on the frequency or intensity of use [15][16][17][18][19][20]. Some emerging studies deploy crowd sourcing survey to effectively collect public opinions (emotions and perceptions) on urban parks and public spaces [21–23]. Some studies also investigated the demographics of park users [24][25][26], and the periods of time parks are used [24].

Regarding park activities, although a considerate number of studies have investigated the level of physical activity in parks [27][25], they were only identified simple events like sedentary, walking, or vigorous. Some studies went beyond this simple categorization and embodied a wider range of park activities [28][29]. However, more studies can still be done on a more fine-grained categorization of activities, as well as on activities driven by organized events as opposed to day-to-day activities such as walking or jogging.

Lastly, it is also worth noting that many past studies on the use of urban parks focused on quantitatively examining the relationship between certain variables and the intensity of use. The independent variables examined include park proximity [15][16], park facilities [15], park quality [30], entrance fees [17], and social demographic characteristics of the neighborhood [15][17].

For the data source and methodology, traditional studies rely heavily on questionnaires and personal interviews. For instance, Schipperijn et al. [8] conducted 14,566 face-to-face interviews with randomly-sampled Danish individuals, and asked these individuals to fill out follow-up questionnaires. Peschardt et al. [31] distributed 686 on-site questionnaires at nine small public urban green spaces to determine how these spaces were used by citizens. Nielsen and Hansen [16] mailed questionnaires to a sample of 2000 adult Danes. Other studies were conducted through direct observations in the parks. For example, many studies, such as the ones by Marquet et al. [20] and Veitch et al. [32], employed the System for Observing Play and Recreation in Communities (SOPARC) [33] to directly observe residents' activities in parks. Similarly, Floyd et al. [25] measured physical activities in parks using a modified version of the System for Observing Play and Leisure Activity in Youth (SOPLAY). Brown et al. [28] used participatory GIS to investigate physical activities in urban parks. Overall, the application of traditional methods to understand park usages and park events is highly time consuming and restrained to smaller areas due to the site-specificity [18].

Recent studies have been incorporating technologies to better understand the use of parks, both through utilizing novel online data sources and more efficient categorization. Commonly used novel data sources include social media data, geo-tracking data from mobile phones, and PPGIS data. For instance, Li et al. [18] retrieved geo-tagged social media check-in records for park visits to examine the frequency of visits. A bivariate correlation analysis was conducted to support the association

between the Weibo check-in data and official visitor statistics, although the strength of correlation ranges from city to city. Larson et al. [19] used geo-tracking data from cell phones to document changes in park visits during the COVID-19 pandemic. Heikinheimo et al. [12] compared four types of data (social media, sports tracking, mobile phone operator and PPGIS data) in a case study of Helsinki, Finland, and examined the ability of these user-generated datasets to provide information on the use of urban parks.

To compare, social media data is highly informative for the leisure time activities being conducted in urban parks [12], but is limited by biases in age groups and the choice to share content publicly [34]; mobile phone data highlights movements [12], but only best represents populations in countries where mobile phones are widely used [35]; PPGIS allows the researcher to ask in-depth questions on park use and preferences [12], but the response rate and its fairness are not guaranteed [13].

For categorization methods, the content analysis of social media data in Heikinheimo's study was done through manual classification of 15,312 Instagram photos and 1,843 Flickr photos. This is again time-consuming and inefficient, and calls for a more automatic method of analyzing social media content on park activities. To compare the best-known commercial image recognition service providers on this task, Ghermandi et al. [29] performed a test using Google Cloud Vision [36], Clarifai [37], and Microsoft Azure Computer Vision [38] to identify human-nature interactions (outdoor recreational activities, biophysical environments, and feelings) in parks. All of these models surpass traditional methods in the efficiency of categorization. However, due to the generic nature of the image recognition services, the tags identified in regards to recreational activities are relatively limited, without sufficient specificity to park-related, event-driven activities. For example, all three services identified people posing for a photograph as the most frequent activity captured in social media imagery. Another precedent to this study is Matasov et al.'s study on COVID-19's impact on the recreational use of Moscow parks, which applied the YOLOv5x neural network to conduct object detection on geo-tagged social media photos.

In conclusion, there are three research gaps in the existing research:

(1) Current studies focus more on the intensity of park usage and level of physical activities (sedentary, walking, vigorous), leaving a gap for more fine-grained studies in the categorization of park events;

(2) For the methodology, traditional studies rely heavily on questionnaires and personal interviews, which is time consuming and restricted;

(3) In recent studies that incorporate technologies, the categorization methods are either inefficient or not specific to park events.

To fill the current research gaps, this study contributes to the literature in these following ways:

(1) By focusing the analysis on the categorization of park events;

(2) By incorporating the use of publicly available imagery to increase the efficiency of analysis;

(3) By proposing transfer learning on pre-trained Convolutional Neural Networks (CNNs) to calibrate the model towards the park event identification task, achieving a 0.876 accuracy and a 0.620 mean average precision.

2. Dataset and Methods

2.1. Research Framework

To more efficiently identify events in urban parks, this research applies Convolutional Neural Networks (CNNs) on images in the *New York City Parks Events Listing* [14] database to conduct multi-label classification of park events. Firstly, we conduct data preprocessing with transfer learning to remove all non-photographic visual media. Secondly, we compare across different machine learning models to determine the best model for the multi-label classification task. See Figure 1.
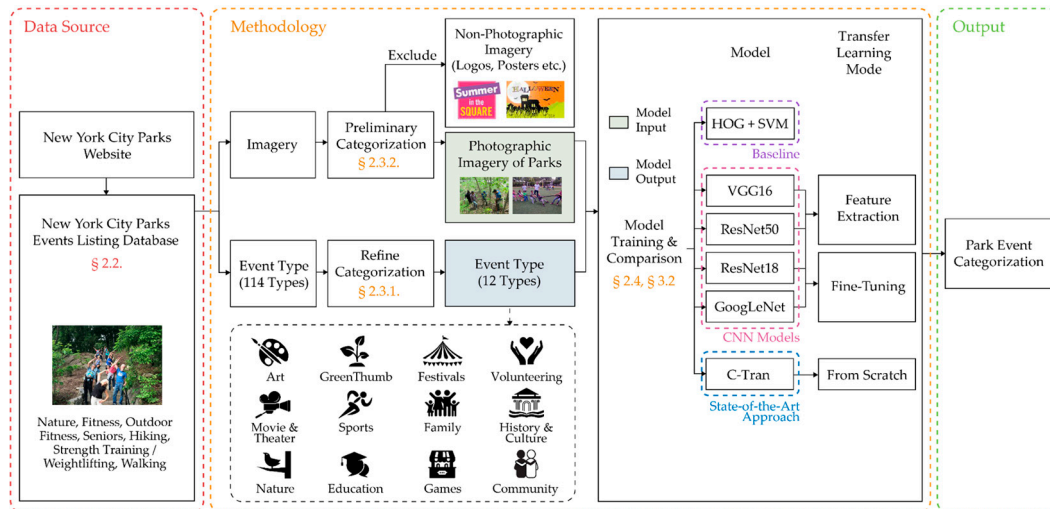
Figure 1. Research Pipeline.

2.2. Dataset

The models are trained on the *New York City Parks Events Listing* database. This database is used to store event information displayed on the *New York City Parks* website, *nyc.gov/parks* (see Figure 2.), which displays events from parks all over New York City. This includes "more than 5,000 individual properties ranging from Coney Island Beach and Central Park to community gardens and Greenstreets" [39]. The *New York City Parks Events Listing* database contains the title, date, time, location, description, contact information, categories, and images of the events since 2013. In total, it contains 11,060 event images, which are linked to 114 event categories. This contains event records from 2013 till August 2, 2021.
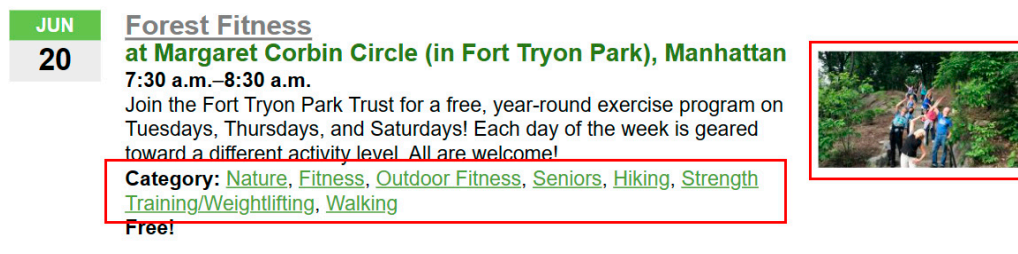

Figure 2. *New York City Parks* website event listing.

2.3. Data Preprocessing

For the purpose of this study, we are only extracting the images and event category information from the dataset, using the Event IDs to link the two together. There are two issues with the original dataset: different levels of specificity in the event categories and the inclusion of non-photographic imagery (logos, posters etc.). Preprocessing was performed to further refine the categorization, reduce the noise, and increase generalizability.

2.3.1. Refining the Categorization

The first issue with the dataset is that the 114 different categories of events in the dataset have different levels of specificity. Some categories are very general, such as "Nature", "Art" or "Volunteer". Other categories are as specific as "Brooklyn Beach Sports Festival" or "MillionTreesNYC: Volunteer: Tree Stewardship and Care". During preprocessing, we manually grouped these categories into larger groups and formed 12 new categories. See Table 1.

Table 1. Event Categorization.

Final Category	Original Category
Art	Art, Arts & Crafts, Art in the Parks: Celebrating 50 Years, Art in the Parks: UNIQLO Park Expressions Grant
GreenThumb	GreenThumb Events, GreenThumb Partner Events, GreenThumb 40th Anniversary, GreenThumb Workshops
Festivals	Festivals, Historic House Trust Festival, Valentine's Day, Halloween, Saint Patrick's Day, Earth Day & Arbor Day, Mother's Day, Father's Day, Holiday Lightings, Santa's Coming to Town, Lunar New Year, Pumpkin Fest, Summer Solstice Celebrations, Easter, Fall Festivals, New Year's Eve, Winter Holidays, Thanksgiving, National Night Out, Black History Month, Women's History Month, LGBTQ Pride Month, Hispanic Heritage Month, Native American Heritage Month, Fourth of July, City of Water Day, She's On Point
Volunteering	Volunteer, MillionTreesNYC: Volunteer: Tree Stewardship and Care, Martin Luther King Jr. Day of Service, MillionTreesNYC: Volunteer: Tree Planting
Film	Film, Free Summer Movies, Theater, Free Summer Theater, Movies Under the Stars, Concerts, Free Summer Concerts, SummerStage, CityParks PuppetMobile
Sports	Fitness, Outdoor Fitness, Running, Bike Month NYC, Hiking, Learn To Ride, Sports, Kayaking and Canoeing, National Trails Day, Brooklyn Beach Sports Festival, Summer Sports Experience, Fishing, Girls and Women in Sports, Bocce Tournament
Family	Best for Kids, Kids Week, CityParks Kids Arts, School Break, Family Camping, Dogs, Dogs in Parks: Town Hall, Seniors, Accessible
History & Culture	History, Historic House Trust Sites, Arts, Culture & Fun Series, Shakespeare in the Parks
Nature	Nature, Birding, Wildlife, Wildflower Week, Cherry Blossom Festivals, Waterfront, Rockaway Beach, Bronx River Greenway, Fall Foliage, Summer on the Hudson, Living With Deer in New York City, Tours, Freshkills Tours, Freshkills Park, Urban Park Rangers, Reforestation Stewardship
Education	Talks, Education, Astronomy, Partnerships for Parks Tree Workshops
Games	Dance, Games, Recreation Center Open House, NYC Parks Senior Games, Mobile Recreation Van Event
Community	Open House New York, Community Input Meetings, Fort Tryon Park Trust, Poe Park Visitor Center, Shape Up New York, City Parks Foundation, Forest Park Trust, City Parks Foundation Adults, Partnerships for Parks Training and Grant Deadlines, Community Parks Initiative, Anchor Parks, Markets, Food

2.3.2. Remove Non-Photographic Imagery

The second issue with the dataset is that it is a mix of photos taken at the parks, and non-photographic visual media such as posters of events and logos of host organizations. To resolve this issue, we introduced feature extraction transfer learning during preprocessing to conduct binary classification and remove the non-photographic images. We applied a VGG16 [40] model pre-trained on the ImageNet [41] dataset, freezing its base layer weights and adding a custom sigmoid layer on top to conduct binary classification. After the top layer was trained on 640 manually-labeled images from the dataset for 25 epochs, with an Adam optimizer and a learning rate of 0.0003, the model achieved a 0.88 training accuracy and a 0.92 accuracy on 160 labeled test images. With this highly

accurate model, we can apply it on the entire dataset to filter out non-photographic images as predicted. This reduced the dataset size from 11,060 images to 7,427 photos.

2.4. Classification Modeling

2.4.1. Model Selection

A wide range of machine learning models were examined in this study to determine the best model for this task, where the inputs are event images and the expected outputs are predictions of the categories of the event.

1. **Baseline:** Histogram of Oriented Gradients (HOG) – Support Vector Machine (SVM) based model

A Histogram of Oriented Gradients (HOG) feature is a feature descriptor used in computer vision and image processing for object detection [42]. The Support Vector Machine (SVM) is a supervised learning algorithm commonly used for classification tasks [43]. A combination of HOG and SVM are incorporated in this study as an example of a traditional approach, where HOG features are extracted from the images and classified through the SVM.

2. Convolutional Neural Networks (CNNs) based models

Convolutional Neural Network (CNN) is a class of artificial neural networks most commonly applied to analyze visual imagery [44]. This study incorporated a selected range of classic CNN models such as VGG16 [40], ResNet50 [45], ResNet18 [45] and GoogLeNet [46]. For each of these CNN models, custom layers including an average pooling layer, a dense layer of 32 neurons (ReLU activation), and a dense layer of 12 neurons (sigmoid activation) were incorporated on top to conduct multi-label classification. The sigmoid layer replaces the conventional softmax layer to accommodate the presence of multiple labels per input image (a park could be used for both fitness and birdwatching). Softmax gives a probability distribution over the entire span of classes, where the 12 probabilities for 12 classes add up to one. By using sigmoid instead, we give each class a number between 1 and 0, and the probabilities do not have to add up to one. Thus, the probability of picking one class is independent from other classes, and we may have multiple labels.

3. State-of-the-Art Approach: C-Tran

C-Tran [47] is a recently proposed model by Lanchantin et al. in 2021, which utilizes Transformers for multi-label image classification. In this study, C-Tran is included as an exemplar of the latest approaches in solving the multi-label classification problem. However, there are limitations to the application of C-Tran in our study due to the discrepancy between the full-image categorization nature of our dataset and the specific dataset assumptions of C-Tran. This is further detailed in Section 2.4.2.

2.4.2. Training

The training process was conducted in the Google Colab environment, using TensorFlow 2.12.0 and a V100 GPU. The models were trained on 80% of the images, with the rest 20% retained for validation and model assessment. Hyperparameters for all CNN models were generally determined through tuning on the VGG16 model, which generated a group of optimized values (batch size = 64, learning rate = 0.0002, number of epochs = 80). These hyperparameters for certain models were slightly tuned in later training. See Table 2. For example, ResNet18 with a batch size of 64 generated suboptimal results. A test of 10 epochs was conducted among ResNet18 models being fine-tuned with batch sizes of relatively 64, 32 and 16, which determined that 32 was the most optimized. All CNN models and C-Tran used the Adam optimizer.

In this study, transfer learning was particularly chosen due to its advantages in efficiency and performance. Training deep neural networks from scratch would require significant computational resources and might not leverage the rich feature-learning already established in networks trained on datasets like ImageNet. Given the specific context of our park events dataset, which is much

smaller and more specialized than vast datasets like ImageNet, it was essential to capitalize on the foundational features such networks have already discerned, like textures or shapes that might be common in park images. Initializing our models with weights from a network pre-trained on ImageNet not only accelerates the training process but also helps in achieving better convergence. Additionally, using transfer learning mitigates the risk of overfitting, especially crucial when working with limited datasets. Accordingly, for each of the CNN models, both feature extraction and fine-tuning techniques were employed for testing. Feature extraction involves freezing the pretrained base layer weights during training, while in fine-tuning all layers are made trainable. The performances of these techniques were then compared to discern the optimal approach for our dataset.

For the C-Tran model, the event description feature from the event listing database was extracted as the image caption for the event image, which should be noted as a limited approach, as the algorithm was originally designed assuming the caption to be a clear and concise description of the image content.

Table 2. Hyperparameters for model training.

| Model	Transfer Learning Mode	Batch Size	Learning Rate	Epochs
VGG16	Feature Extraction	64	0.0002	80
	Fine-Tuning	64	0.0002	80
ResNet50	Feature Extraction	64	0.0002	100
	Fine-Tuning	64	0.0002	70
ResNet18	Feature Extraction	32	0.0002	20
	Fine-Tuning	32	0.0001	10
GoogLeNet	Feature Extraction	64	0.0002	80
	Fine-Tuning	64	0.0002	60
C-Tran	From Scratch	1	0.00001	40

2.4.3. Evaluation Metrics

This study incorporates both the accuracy and the mean Average Precision (mAP) metrics to evaluate the model performance. In calculation of the accuracy, we treat the classification of each model as an independent task, and calculate the average accuracy across labels. We also incorporated the mAP, a commonly used metric to evaluate object detection models, as it is a relatively comprehensive evaluation metric that takes into account both precision and recall for each class or label.

3. Results

3.1. Descriptive Statistics

Figure 3 shows the distribution of images across different labels in the dataset after non-photographic imagery was removed (as described in Section 2.3.2.). 'Family', 'Nature', and 'Film' are the three categories that occurred most frequently. 'GreenThumb' and 'Volunteer' only contain a very small number of images.
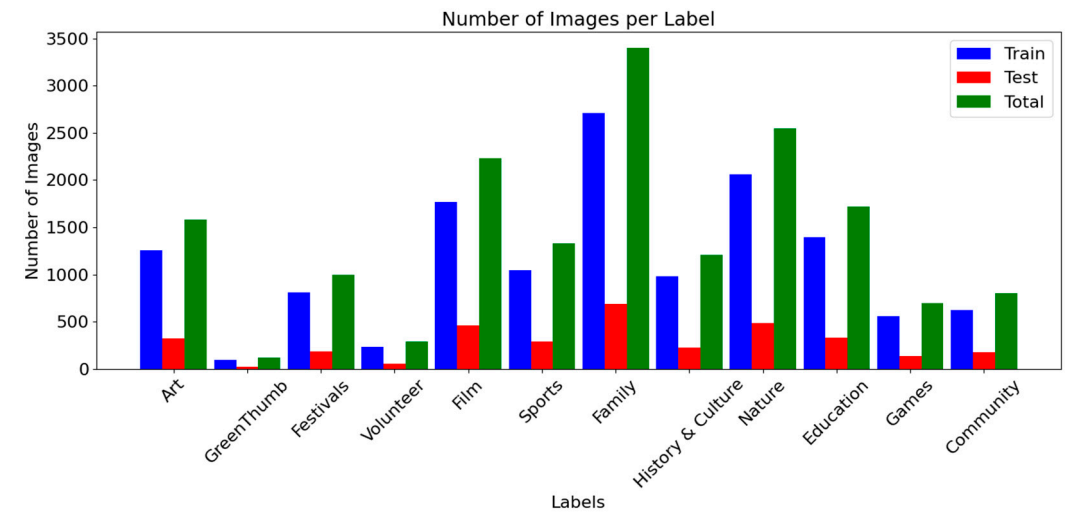
Figure 3. Image distribution across event types.

Figure 4 illustrates the distribution of selected event categories within New York City parks. Events categorized under 'Film' are prevalent across numerous locations, suggesting that many of these parks are equipped for outdoor film screenings or theatrical performances. Conversely, while the 'Art' category displays a peak value of 249 events at a single park, such events are less widespread. This limited distribution indicates that specialized facilities are needed for art events, possibly making them less accessible to residents citywide. In a similar vein, parks featuring 'Nature' events are predominantly located towards the city's outskirts, which aligns with expectations. Figure A1 presents the figures for the rest of the event categories.

(**a**) Art.   (**b**) Film.

I Sports.   (**d**) Nature.
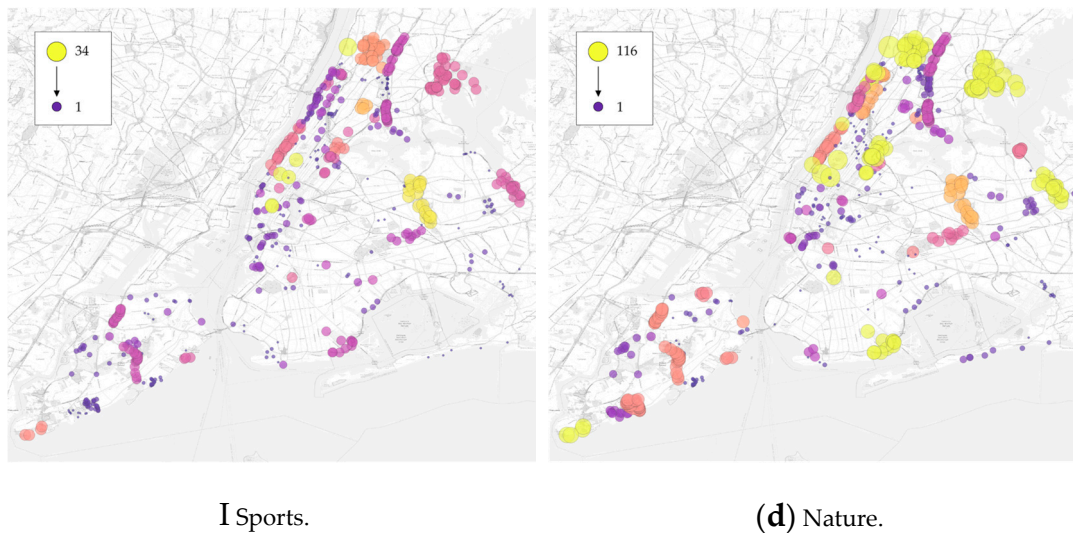
Figure 4. Distribution of selected park event categories across parks in New York City.

In the diverse urban tapestry of New York City, parks emerge as dynamic spaces of community interaction and learning. Figure 5 shows the co-occurrence matrix of different event types. We observed that 'Family' and 'Art', 'Family' and 'Film', 'Family' and 'Nature', 'Family' and 'Education', and 'Nature' and 'Education' are frequent co-occurrences. The co-occurrence of events such 's 'Family & 'rt' underscores the c'ty's commitment to fostering a vibrant arts culture, making it accessible to audiences of all ages. Outdoor movie sessions, exemplified by t'e 'Family & F'lm' pairing, showcase the pa'ks' ability to transform into open-air theaters, creating unique urban experiences. The conjunction 'f 'Family & Nat're' a'd 'Family & Educat'on' emphasizes the pa'ks' role as both recreational escapes and vital educational hubs. Parks not only offer families a chance to reconnect with nature but also provide hands-on educational experiences. Lastly, the overlap betwe'n 'Nature & Educat'on' reiterates the importance of these urban green spaces in fostering environmental awareness and stewardship among its citizens. Such multifaceted interactions in New York City parks highlight their indispensable role in enhancing the city's cultural, recreational, and educational landscape.
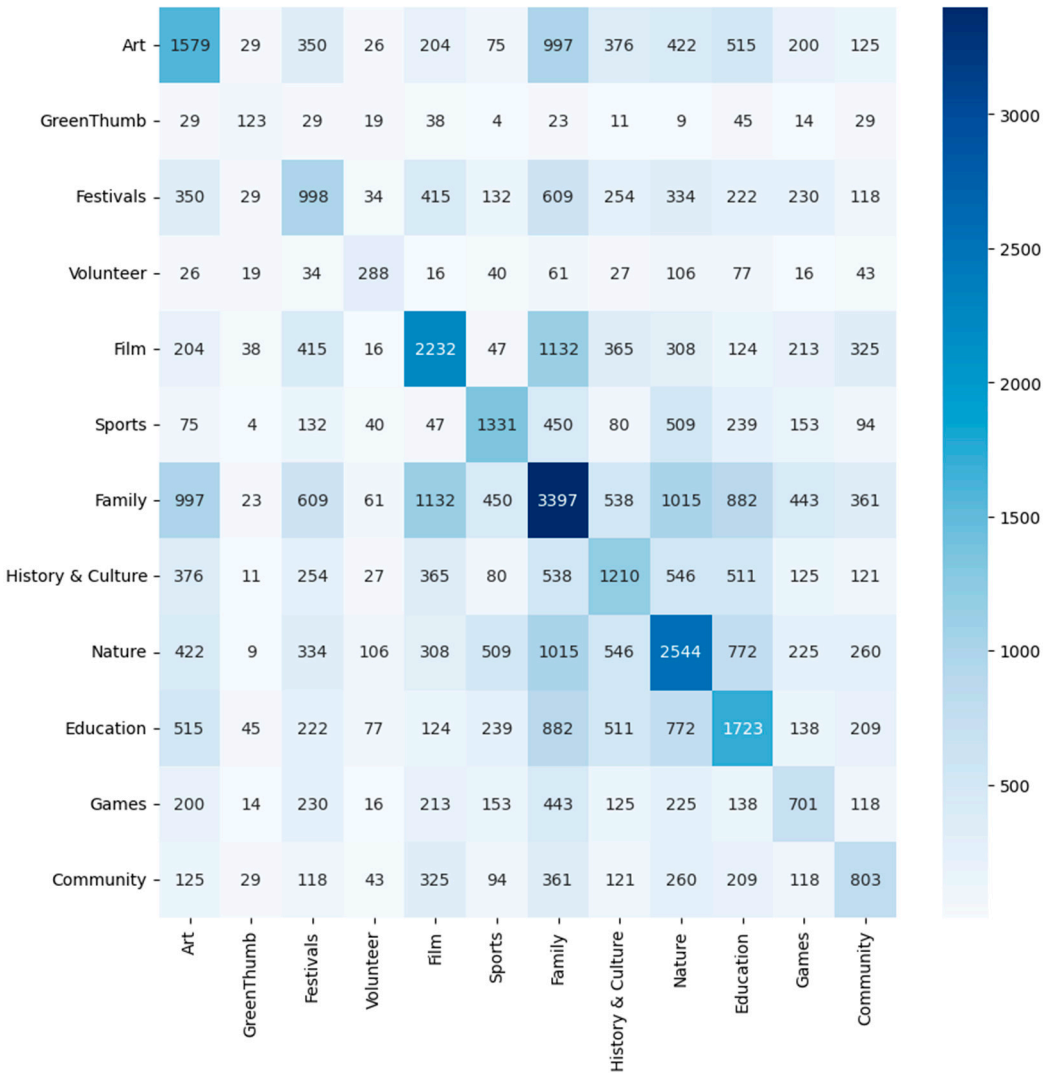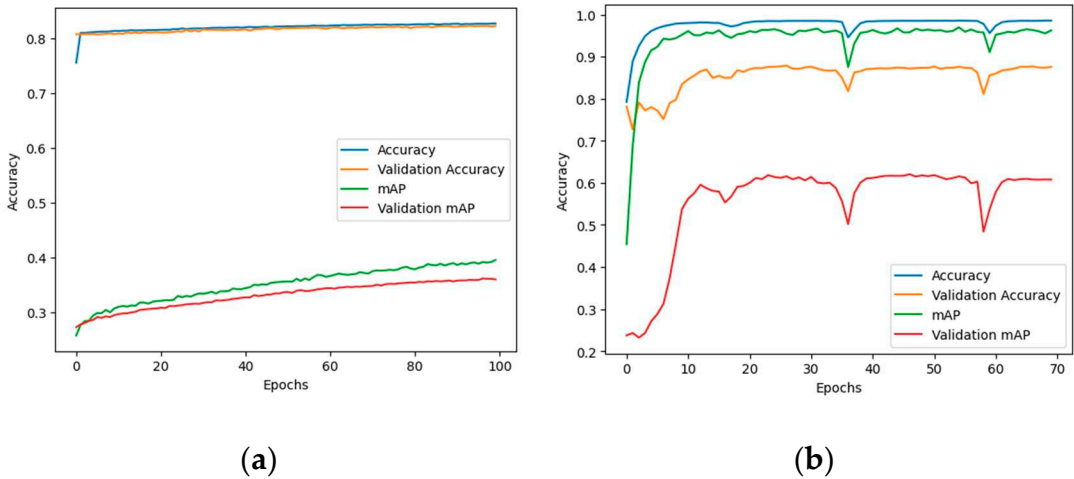
Figure 5. Event type co-occurrence matrix.

3.2. Overall Performance of Event Classification

Figure 6 presents the accuracy and mean Average Precision change throughout the training process for both feature extraction and fine-tuning on ResNet50, as an example comparison for these two transfer learning approaches.

(a)

(b)

Figure 6. Comparing accuracy and mean Average Precision between different transfer learning approaches. (**a**) ResNet50 Feature Extraction. (**b**) ResNet50 Fine-Tuning.

Table 3 presents all results from the models examined, including the baseline HOG + SVM approach and the state-of-the-art C-Tran model. Among all the examined approaches, fine-tuning on the ResNet50 model achieved the best performance in both accuracy and mean Average Precision, outperforming ResNet18 and GoogLeNet (InceptionV3) by a small margin. This suggests that ResNet50 was the most capable in capturing the features that indicate park events and recreational human activities in this dataset.

Table 3. Validation Accuracy and mAP.

Model	Transfer Learning Mode	Accuracy	mAP *
HOG + SVM	From Scratch	0.861	0.345
VGG16	Feature Extraction	0.844	0.462
	Fine-Tuning	0.854	0.564
ResNet50	Feature Extraction	0.823	0.360
	Fine-Tuning	**0.876**	**0.620**
ResNet18	Feature Extraction	0.809	0.291
	Fine-Tuning	0.870	0.601
GoogLeNet	Feature Extraction	0.857	0.551
	Fine-Tuning	0.876	0.602
CTran	From Scratch	-	0.200

* mean Average Precision.

Figure 7 presents the normalized confusion matrices for each label, where the x axis is the prediction (with a threshold of 0.5) and the y axis is the ground truth. These graphs show that for all labels, true negatives compose the majority of the confusion matrices, and false positives compose the least percentage. This suggests that the model is generally conservative in its predictions. There are missed opportunities in the labels 'GreenThumb', 'Festivals', 'Volunteer', 'History & Culture', 'Education', 'Games', and 'Community', where false negatives outnumber true positives. Among these labels, 'GreenThumb' (99) and 'Volunteer' (233) are labels with a very low portion of corresponding training images. 'Festivals' (809), 'History & Culture' (984), and 'Education' (1,393) are labels with relatively sufficient training images, but still exhibit a concerning number of false negatives, which suggests that the mo'el's inability to accurately predict these categories is potentially due to other factors such as data quality and label ambiguity. 'Games' (560) and 'Community' (625) are labels with a medium number of images, and the cause of underperformance is hard to determine. The model is particularly successful in predicting the presence of 'Film', 'Family', and 'Nature'. These are also the three categories that compose the overwhelming majority of the training dataset, with each category containing more than 1,700 images.

Another contributing factor for the accurate identification of events under t'e 'F'lm', 'Fam'ly', a'd 'Nat're' categories could be the distinct features found within the parks themselves. These unique amenities or landmarks may be intrinsically tied to the events in these categories. For instance, parks hosti'g 'F'lm' events may have dedicated open spaces or amphitheaters suitable for large audiences, those emphasizi'g 'Fam'ly' events might possess playgrounds or picnic areas designed for family gatherings, and parks with freque't 'Nat're' events could be characterized by trails, water bodies, or other natural landmarks. Such distinct features could make categorizing events in these parks more straightforward.
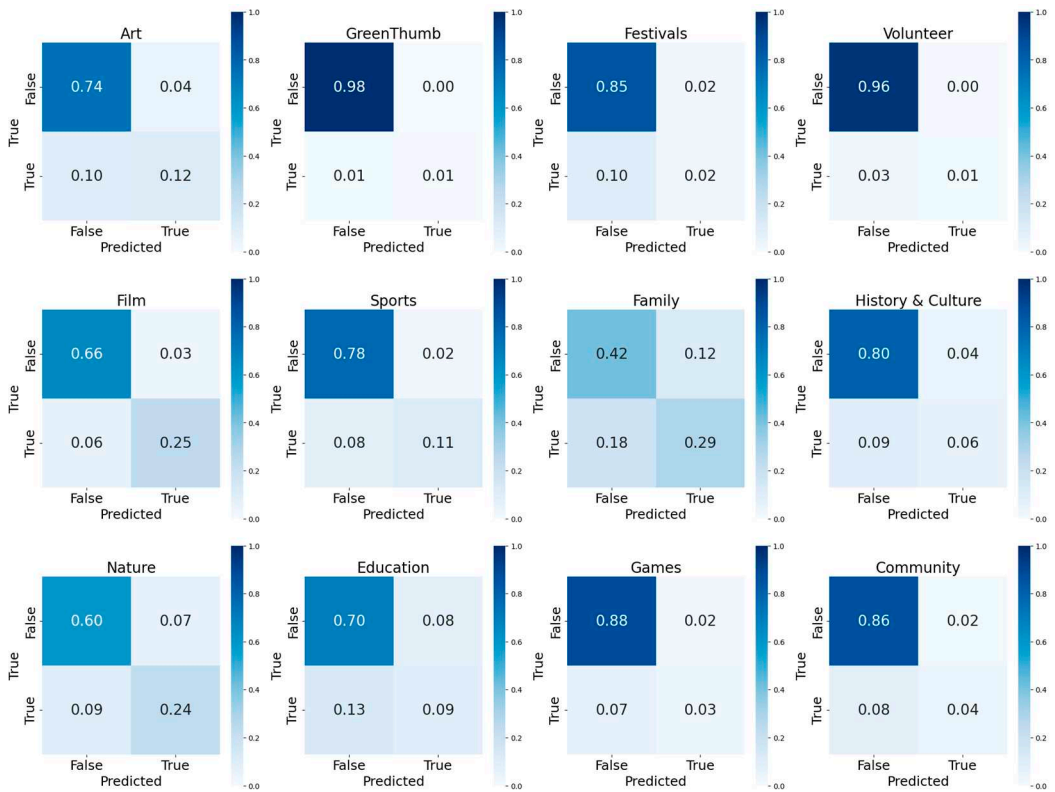
Figure 7. Normalized confusion matrices (X axis = predicted classes; Y axis = true classes).

Figure 8 presents the normalized co-occurrence matrix for the true and predicted labels, where the x axis represents the predicted classes, and the y axis represents the true classes. On the diagonal, 'Film', 'Sports', 'Nature' and 'Family' are the four labels with the highest percentage of successful classification. 'Festivals' is a label that the model specifically struggles with. It is also worth noting that, due to the multi-label nature of the classification task, the ideal for this matrix is not necessarily to have high values only along the diagonal. For example, high values occur on the intersections of the 'Family' row and the 'Art', 'Film', 'Nature' and 'Education' columns. This is exactly in correspondence with what we observed in Section 3.1. about the co-occurrences in the dataset, potentially suggesting that the model was successful in identifying genuine patterns from the data.
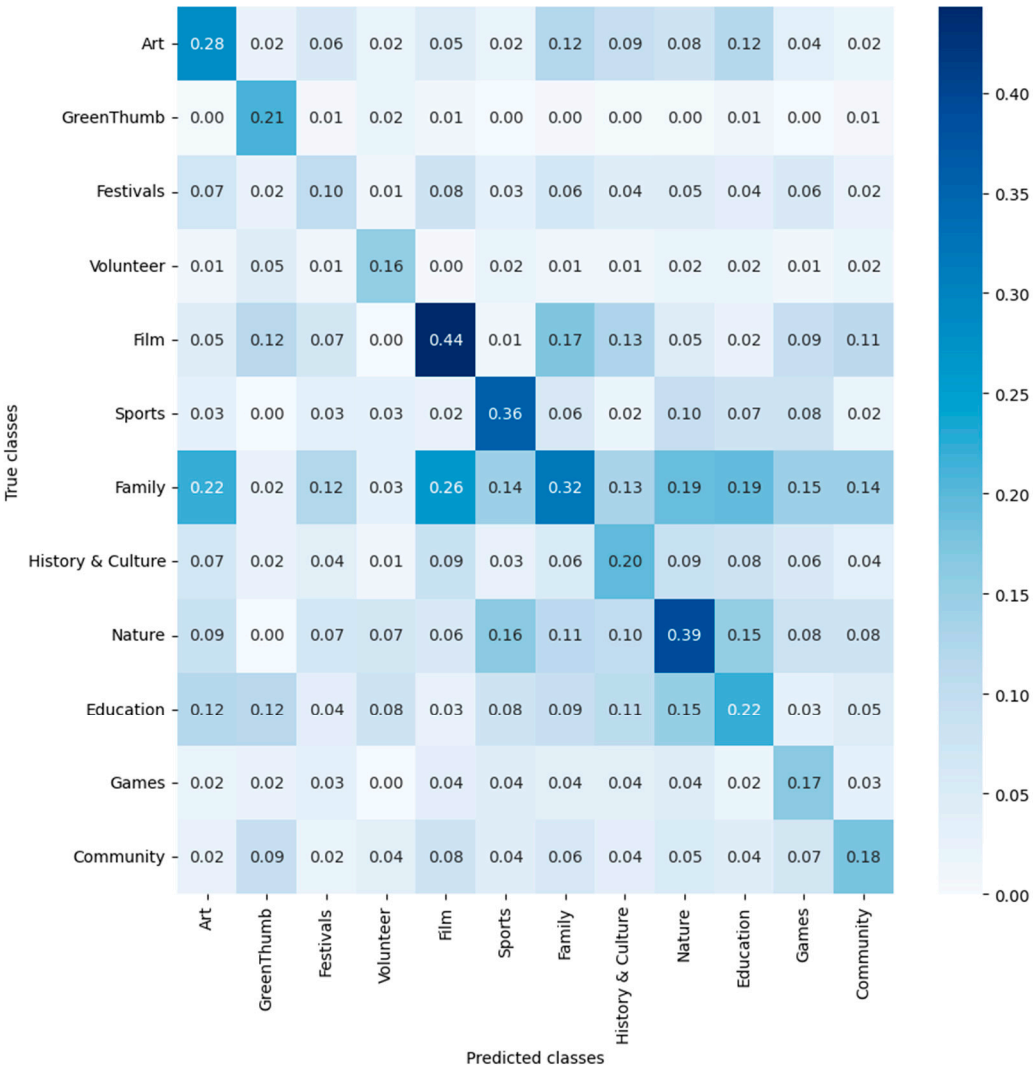
Figure 8. Event type co-occurrence matrix (X axis = predicted classes; Y axis = true classes).

3.3. Transfer Learning Approaches

It is worth noting that for this task, fine-tuning on all CNN models outperforms feature extraction transfer learning, and some models such as ResNet18 even showed significant performance differences. This might suggest that there is a limited similarity between the task of the pre-trained model (object recognition based on ImageNet) and the target domains of this task. This can be attributed to the nature of the dataset where in a lot of the images, the model needs to recognize the gesture of the human(s) to determine the label, while the ImageNet dataset is organized only around nouns [36]. Another thing this suggests is the complexity of the task, since the situation could indicate that the t′sk's complexity exceeds what can be adequately addressed by the feature extraction approach. Fine-tuning, on the other hand, allows the model to adapt to the specific features of the New York City park event images. Figure 9 presents examples of park event images, their true labels and predicted labels. This offers a tangible representation of the mo′el's predictive capabilities, showcasing instances where the model successfully identified the event type as well as moments of misclassification. By observing the images side-by-side with their labels, readers can gain insights into the nuanced features the model potentially considers when making its predictions.

(**a**) True: {*Art, Family*}; Predicted: {*Art, Family, Education*}.
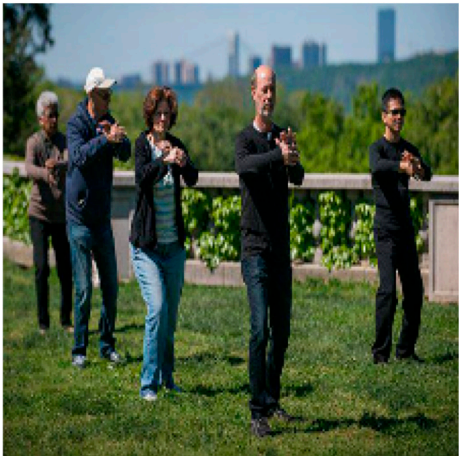
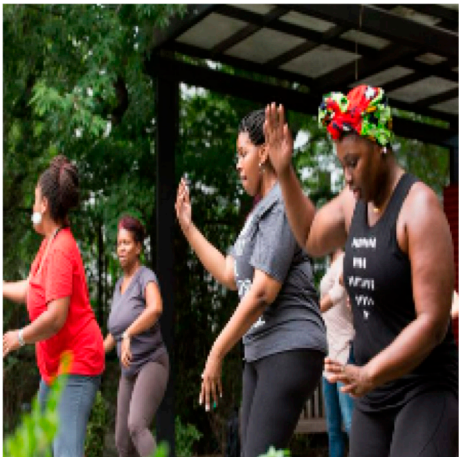(**b**) True: {*Sports*}; Predicted: {*Sports*}.

(**c**) True: {*GreenThumb, Volunteer, Education*}; Predicted: {*GreenThumb, Volunteer, Education*}.

(**d**) True: {*Nature, History & Culture*}; Predicted: {*Nature, Family, Volunteer*}.

(**e**) True: {*Art, Festivals, Film, History & Culture*}; Predicted: {*Art, Festivals, Film, History & Culture*}.

(**f**) True: {*Art, Sports, Games*}; Predicted: {*Sports, Games*}.

Figure 9. Example images, their true labels and predictions from ResNet50 Fine-Tunin**4.**

4. Conclusion

Understanding park events and being able to categorize them is crucial to understanding parks and their role in urban areas. This study uses the images and event category information in the *New York City Parks Events Listing* database to train a Convolutional Neural Network that categorizes park events represented in images. Upon evaluating various models, it was determined that ResNet50 emerged as the most proficient in the event categorization task, achieving an accuracy of 0.876 and a mAP of 0.620, outperforming the other models compared. The results demonstrate the potential of deep learning techniques in automating the categorization process of park events, which can provide invaluable insights into the activities and cultural dynamics within urban parks. The trained CNN exhibited promise in recognizing and differentiating between various event types, highlighting the diverse range of activities that urban parks can host. Furthermore, accurate categorization can aid city planners and park administrators in making informed decisions about resource allocation, event scheduling, and infrastructure development tailored to the unique needs of different event types. As urban areas continue to grow and evolve, leveraging technology to better understand and optimize the use of public spaces like parks becomes increasingly vital.

Future avenues of research encompass both the application of our trained model to unlabeled datasets and the expansion of our labeled datasets to further hone the model's accuracy. To begin with, our model can be deployed on unlabeled datasets from popular social media platforms like Instagram and Flickr. This would enable efficient categorization of park-related event images, providing deeper insights into event distributions and enhancing our understanding of the diverse roles urban parks play within communities. Furthermore, integrating more labeled data, sourced from similar park event listing websites such as the one from Millennium Park in Chicago [48], can bolster the model's performance, ensuring more accurate and robust categorizations in future applications. To maintain consistent model performance on the expanded dataset, additional experiments may be required. These will focus on identifying the best model for the preprocessing step detailed in Section 2.3.2. The goal is to achieve a performance comparable to the current preprocessing task, which boasts an accuracy of approximately 0.92.

Appendix A

(**a**) GreenThumb.

(**b**) Festivals.
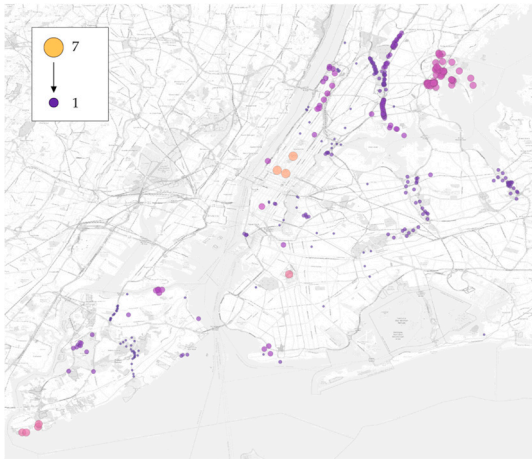
(**c**) Volunteering.
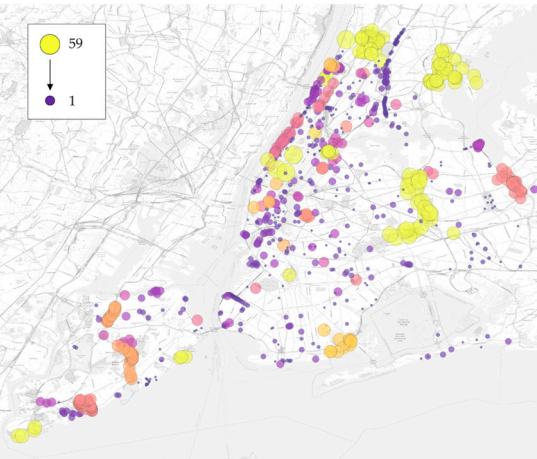
(**d**) Family.
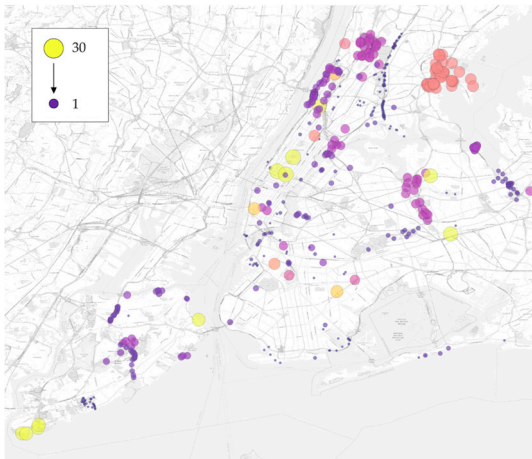
(**e**) History & Culture.

(**f**) Education.

(**g**) Games.   (**h**) Community.
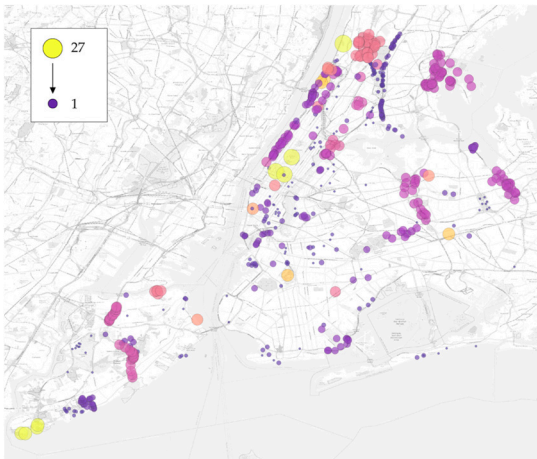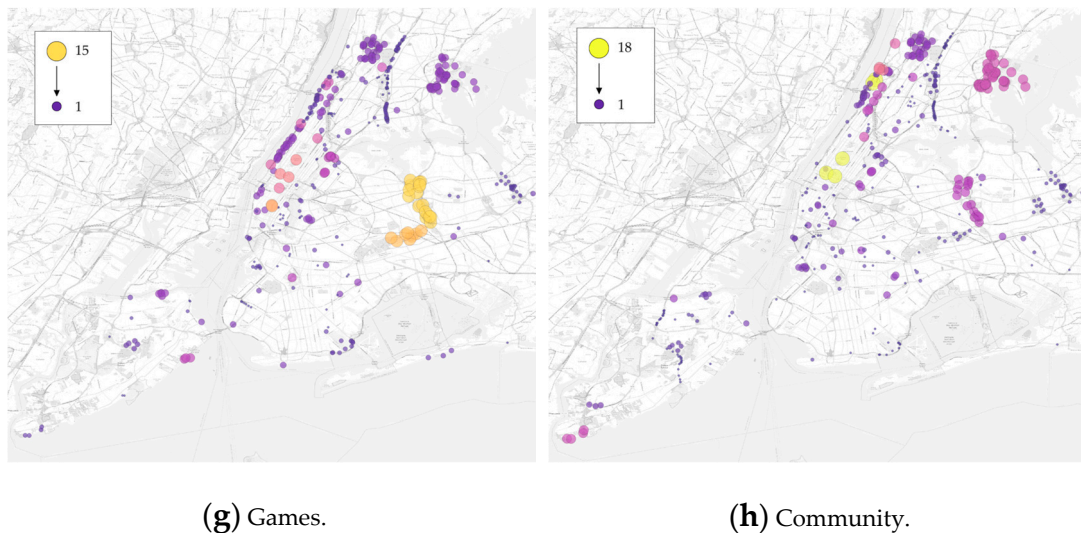
Figure A1. Distribution of other park event categories across parks in New York City.

References

1. Konijnendijk, C.; Annerstedt, M.; Nielsen, A.B.; Maruthaveeran, S. *Benefits of Urban Parks: A Systematic Review*; International Federation of Parks and Recreation Administration, 2013;

2. Sadeghian, M.M.; Vardanyan, Z. The Benefits of Urban Parks, a Review of UrbanResearch. *J Nov. Appl Sci.* **2013**, *2 (8)*, 231–237.

2. Smith, A.; Vodicka, G. *Events in London's Parks: The Friends' Perspective*; Zenodo, 2020;

3. Smith, A.; Vodicka, G.; Colombo, A.; Lindstrom, K.N.; McGillivray, D.; Quinn, B. Staging City Events in Public Spaces: An Urban Design Perspective. *IJEFM* **2021**, *12*, 224–239, doi:10.1108/IJEFM-10-2020-0063.

4. Smith, A.; Osborn, G.; Vodicka, G. Private Events in a Public Park: Contested Music Festivals and Environmental Justice in Finsbury Park, London. In *Whose Green City?*; Plüschke-Altof, B., Sooväli-Sepping, H., Eds.; Sustainable Development Goals Series; Springer International Publishing: Cham, 2022; pp. 83–102 ISBN 978-3-031-04635-3.

5. Neal, S.; Bennett, K.; Jones, H.; Cochrane, A.; Mohan, G. Multiculture and Public Parks: Researching Super-Diversity and Attachment in Public Green Space: Multiculture and Public Parks. *Popul. Space Place* **2015**, *21*, 463–475, doi:10.1002/psp.1910.

6. Citroni, S.; Karrholm, M. Neighbourhood Events and the Visibilisation of Everyday Life: The Cases of Turro (Milan) and Norra Fäladen (Lund). *European Urban and Regional Studies* **2019**, *26*, 50–64, doi:10.1177/0969776417719489.

7. Schipperijn, J. et al. 2010. Factors influencing the use of green space: Results from a Danish national representative survey. *Landscape and Urban Planning* **2010**, *95(3)*, 130–137, doi:10.1016/j.landurbplan.2009.12.010.

8. Moran, M.R.; Rodríguez, D.A.; Cotinez-O'Ryan, A.; Miranda, J.J. Park Use, Perceived Park Proximity, and Neighborhood Characteristics: Evidence from 11 Cities in Latin America. *Cities* **2020**, *105, 102817*, doi:10.1016/j.cities.2020.102817.

9. Neuvonen, M.; Sievänen, T.; Tönnes, S.; Koskela, T. Access to Green Areas and the Frequency of Visits – a Case Study in Helsinki. *Urban Forestry & Urban Greening*, **2007**, *6(4)*, 235–47, doi:10.1016/j.ufug.2007.05.003.

10. *Analysis of Activities and Participation Questionnaire*; South Gloucestershire Council: Page Park.

11. Heikinheimo, V.; Tenkanen, H.; Bergroth, C.; Järv, O.; Hiippala, T.; Toivonen, T. Understanding the Use of Urban Green Spaces from User-Generated Geographic Information. *Landscape and Urban Planning* **2020**, *201(103845)*, doi:10.1016/j.landurbplan.2020.103845.

12. Brown, G. A Review of Sampling Effects and Response Bias in Internet Participatory Mapping (PPGIS/PGIS/VGI): Sampling Effects and Response Bias in Internet Participatory Mapping. *Trans. in GIS* **2017**, *21*, 39–56, doi:10.1111/tgis.12207.

13. NYC parks events listing – event listing: NYC open data. Available online: https://data.cityofnewyork.us/browse?Data-Collection_Data-Collection=NYC+Parks+Events&sortBy=alpha (accessed on 9 December 2021)

14. Kaczynski, A.T.; Besenyi, G.M.; Stanis, S.A.; Koohsari, M.J.; Oestman, K.B.; Bergstrom, R.; Potwarka, L.R.; Reis, R.S. Are Park Proximity and Park Features Related to Park Use and Park-Based Physical Activity among Adults? Variations by Multiple Socio-Demographic Characteristics. *International Journal of Behavioral Nutrition and Physical Activity* **2014**, *11(1)*, doi:10.1186/s12966-014-0146-4.

15.   Nielsen, T.S.; Hansen, K.B. Do Green Areas Affect Health? Results from a Danish Survey on the Use of Green Areas and Health Indicators. *Health & Place* **2007**, *13(4)*, 839–50, doi:10.1016/j.healthplace.2007.02.001.

16.   Bjork, J.; Albin, M.; Grahn, P.; Jacobsson, H.; Ardo, J.; Wadbro, J.; Ostergren, P.O.; Skarback, E. Recreational Values of the Natural Environment in Relation to Neighbourhood Satisfaction, Physical Activity, Obesity and Wellbeing. *Journal of Epidemiology & Community Health* **2008**, *62(4)*. https://doi.org/10.1136/jech.2007.062414.

17.   Larson, L.R.; Zhang, Z.; Oh, J.I.; Beam, W.; Ogletree, S.S.; Bocarro, J.N.; Lee, K.J. et al. Urban Park Use during the COVID-19 Pandemic: Are Socially Vulnerable Communities Disproportionately Impacted? *Frontiers in Sustainable Cities* **2021**, *3*, doi:10.3389/frsc.2021.710243.

18.   Li, F.; Li, F.; Li, S.; Long, Y. Deciphering the Recreational Use of Urban Parks: Experiments Using Multi-Source Big Data for All Chinese Cities. *Science of The Total Environment* **2020**, *701(134896)*, doi:10.1016/j.scitotenv.2019.134896.

19.   Dong, L.; Jiang, H.; Li, W.; Qiu, B.; Wang, H.; Qiu, W. Assessing Impacts of Objective Features and Subjective Perceptions of Street Environment on Running Amount: A Case Study of Boston. *Landscape and Urban Planning* **2023**, *235*, 104756, doi:10.1016/j.landurbplan.2023.104756.

20.   Su, N.; Li, W.; Qiu, W. Measuring the Associations between Eye-Level Urban Design Quality and on-Street Crime Density around New York Subway Entrances. *Habitat International* **2023**, *131*, 102728, doi:10.1016/j.habitatint.2022.102728.

21.   Qiu, W.; Zhang, Z.; Liu, X.; Li, W.; Li, X.; Xu, X.; Huang, X. Subjective or Objective Measures of Street Environment, Which Are More Effective in Explaining Housing Prices? *Landscape and Urban Planning* **2022**, *221*, 104358, doi:10.1016/j.landurbplan.2022.104358.

22.   Qiu, W.; Li, W.; Liu, X.; Zhang, Z.; Li, X.; Huang, X. Subjective and Objective Measures of Streetscape Perceptions: Relationships with Property Value in Shanghai. *Cities* **2023**, *132*, 104037, doi:10.1016/j.cities.2022.104037.

23.   Kaczynski, A.T.; Potwarka, L.R.; Smale, B.J.; Havitz, M.E. Association of Parkland Proximity with Neighborhood and Park-Based Physical Activity: Variations by Gender and Age. *Leisure Sciences* **2009**, 31(2), 174–91, doi:10.1080/01490400802686045.

24.   Floyd, M.F.; Spengler, J.O.; Maddock, J.E.; Gobster, P.H.; Suau, L,J. Park-Based Physical Activity in Diverse Communities of Two U.S. Cities: An Observational Study. *American Journal of Preventive Medicine* **2008**, *34(4)*, 299-305, ISSN 0749-3797, doi:10.1016/j.amepre.2008.01.009.

25.   Lin, B.B.; Fuller, R.A.; Bush, R.; Gaston, K.J.; Shanahan, D.F. Opportunity or Orientation? Who Uses Urban Parks and Why. *PLoS ONE* **2014**, *9*, e87422, doi:10.1371/journal.pone.0087422.

26.   Evenson, K.R.; Jones, S.A.; Holliday, K.M.; Cohen, D.A.; McKenzie, T.L. Park Characteristics, Use, and Physical Activity: A Review of Studies Using SOPARC (System for Observing Play and Recreation in Communities). *Preventive Medicine* **2016**, *86*, 153–166, doi:10.1016/j.ypmed.2016.02.029.

27.   Brown, G.; Schebella, M.F.; Weber, D. Using participatory GIS to measure physical activity and urban park benefits. *Landscape and Urban Planning* 2014, *121*, 34-44, ISSN 0169-2046, doi:10.1016/j.landurbplan.2013.09.006.

28.   Ghermandi, A.; Depietri, Y.; Sinclair, M. 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* **2022**, *217*, 104261, doi:10.1016/j.landurbplan.2021.104261.

29.   Coles, R.W.; Bussey, S.C. Urban Forest Landscapes in the UK — progressing the Social Agenda. *Landscape and Urban Planning* **2000**, *52(2–3)*, 181–188, doi:10.1016/s0169-2046(00)00132-8.

30.   Peschardt, K.K.; Schipperijn, J.; Stigsdotter, U.K. Use of Small Public Urban Green Spaces (SPUGS). *Urban Forestry & Urban Greening* **2012**, *11(3)*, 235–44, doi:10.1016/j.ufug.2012.04.002.

31.   Veitch, J.; Ball, K.; Crawford, D.; Abbott, G.R.; Salmon, J. Park Improvements and Park Activity. *American Journal of Preventive Medicine* **2012**, *42*, 616–619, doi:10.1016/j.amepre.2012.02.015.

32.   McKenzie, T.L.; Cohen, D.A.; Sehgal, A.; Williamson, S.; Golinelli, D. System for observing play and recreation in communities (SOPARC): reliability and feasibility measures. *J. Phys. Act. Health* **2006**, *3 (s1)*, S208–S222, doi:10.1123/jpah.3.s1.s208.

33.   Heikinheimo, V.; Minin, E.D.; Tenkanen, H.; Hausmann, A.; Erkkonen, J.; Toivonen, T. User-Generated Geographic Information for Visitor Monitoring in a National Park: A Comparison of Social Media Data and Visitor Survey. *IJGI* **2017**, *6*, 85, doi:10.3390/ijgi6030085.

34.   Ahas, R.; Silm, S.; Saluveer, E.; Järv, O. Modelling Home and Work Locations of Populations Using Passive Mobile Positioning Data. In *Location Based Services and TeleCartography II*; Gartner, G., Rehrl, K., Eds.; Lecture Notes in Geoinformation and Cartography; Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; pp. 301–315 ISBN 978-3-540-87392-1.

35.   Vision AI. Available online: https://cloud.google.com/vision (accessed on 18 August 2023)

36.   General Image Recognition. Available online: https://www.clarifai.com/models/general-image-recognition (accessed on 18 August 2023)

37. Azure AI Vision with OCR and AI. Available online: https://azure.microsoft.com/en-us/products/ai-services/ai-vision#:~:text=Azure%20AI%20Vision%20is%20a,)%2C%20and%20responsible%20facial%20recognition. (accessed on 18 August 2023)

38. About Parks: NYC Parks. Available online: https://www.nycgovparks.org/about (accessed on 18 August 2023)

39. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7-9 May 2015.

40. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. IEEE Computer Vision and Pattern Recognition (CVPR), 2009.

41. Freeman, W.T.; Roth, M. Orientation Histograms for Hand Gesture Recognition. In Proceedings of the IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition, Zurich, Switzerland, June, 1995

42. Cortes, C.; Vapnik, V. Support-vector networks. *Machine Learning*, **1995**, *20 (3)*, 273–297. doi:10.1007/BF00994018.

43. Valueva, M.V.; Nagornov, N.N.; Lyakhov, P.A.; Valuev, G.V.; Chervyakov, N.I. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and Computers in Simulation*, **2020**, *Elsevier BV. 177*, 232–243. doi:10.1016/j.matcom.2020.04.031.

44. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

45. Szegedy, C. et al., Going deeper with convolutions, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

46. Lanchantin, J.; Wang, T.; Ordonez, V.; Qi, Y. General Multi-label Image Classification with Transformers, Nashville, TN, USA, 19-25 June 2021.

47. City of Chicago: Millennium Park Calendar. Available online: https://www.chicago.gov/city/en/depts/dca/supp_info/mp_calendar.html (accessed on 18 August 2023).