The Spirit of the City: Using Social Media to Capture Neighborhood Ambiance

MIRIAM REDI, Nokia Bell Labs, United Kingdom
LUCA MARIA AIELLO, Nokia Bell Labs, United Kingdom
ROSSANO SCHIFANELLA, University of Turin, Italy
DANIELE QUERCIA, Nokia Bell Labs, United Kingdom

Place ambiance has a huge influence over how we perceive places. Despite its importance, ambiance has been crucially overlooked by urban planners and scientists alike, not least because it is difficult to record and analyze at scale. We explored the possibility of using social media data to reliably map the ambiance of neighborhoods in the entire city of London. To this end, we collected geo-referenced picture tags from Flickr and matched those tags with the words in a newly created ambiance dictionary. In so doing, we made four main contributions: i) map the ambiance of London neighborhoods; ii) ascertain that such a mapping meets residents' expectations, which are derived from a survey we conducted; iii) show that computer vision techniques upon geo-referenced pictures are of predictive power for neighborhood ambiance; and iv) explain each prediction of a neighborhood's ambiance by identifying the picture that best reflects the meaning of that ambiance (e.g., artsy) in that neighborhood (e.g., South Kensington—the richest and most traditional neighborhood—and Shoreditch—among the most progressive and hipster neighborhoods in the city—are both 'artsy' but in very different ways). The combination of the predictive power of mapping ambiance from images and the ability to explain those predictions makes it possible to discover hidden gems across the city at an unprecedented scale.

CCS Concepts: • **Human-centered computing** → *Computer supported cooperative work*; *Social media*; • **Computing methodologies** → *Computer vision representations*;

Additional Key Words and Phrases: ambiance; computer vision; London; Flickr

ACM Reference Format:
Miriam Redi, Luca Maria Aiello, Rossano Schifanella, and Daniele Quercia. 2018. The Spirit of the City: Using Social Media to Capture Neighborhood Ambiance. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 144 (November 2018), 18 pages. https://doi.org/10.1145/3274413

1 INTRODUCTION

Nullus locus sine Genio, there is no place without a spirit. For the Romans, *Genius Loci* was the divinity protecting a place. In modern times, *Genius Loci* is the location's distinctive ambiance or character [20]. Urban areas can have different "spirits": they can be creative, lively or, say, cosmopolitan. The spirit, or ambiance, of places impacts our daily perceptions and, in the long run, our lives [10, 11].

Authors' addresses: Miriam Redi, Nokia Bell Labs, Cambridge, United Kingdom, miriam.redi@gmail.com; Luca Maria Aiello, Nokia Bell Labs, Cambridge, United Kingdom, luca.aiello@nokia-bell-labs.com; Rossano Schifanella, University of Turin, Turin, Italy, schifane@di.unito.it; Daniele Quercia, Nokia Bell Labs, Cambridge, United Kingdom, quercia@cantab.net.

2573-0142/2018/11-ART144 $15.00
https://doi.org/10.1145/3274413

Despite the importance of urban ambiance, computer scientists have rarely looked at it from a computational perspective. Existing research mostly focuses on a limited set of ambiance dimensions: it has focused on indoor venue atmosphere [26], socio-demographic neighborhood attributes [28], and subjective perceptions of urban scenes in terms of safety [7] and beauty [23].

That might be because urban ambiance is intangible ("We grasp the atmosphere before we identify its details or understand it intellectually" [21]), multi-faceted ("The outcome of a complex composition of physical, physiological, psychological, sociological and cultural criteria" [8]), and is therefore hard to define and measure. In his 1970's paper "The Experience of Living in Cities" [16], Stanley Milgram aptly wrote: *"It may seem that urban atmosphere is too evanescent a quality to be reduced to a set of measurable variables, but I do not believe the matter can be judged before substantial effort has been made in this direction."*

Today, user-generated content and mobile devices represent an unprecedented opportunity to study place ambiance [30]. In this work, we aimed at facing the "substantial effort" Milgram was referring to, and we did so by designing a new way of automatically measuring urban ambiance. This way relies on the application of computer vision on social media pictures. We designed an end-to-end framework for urban ambiance definition, mapping, prediction and understanding, and evaluated its effectiveness with qualitative and quantitative studies. More specifically, this work makes four main contributions:

- *Literature-driven Ambiance Taxonomy.* For the first time, we defined a vocabulary of terms related to urban ambiance (e.g., "gloomy", "peaceful") and structured them in a 3-level taxonomy (§3).
- *Ambiance Mapping based on social media.* We matched the ambiance terms with tags on Flickr pictures that were geo-referenced in London, aggregated the resulting ambiance tags at neighborhood level, and accordingly scored all neighborhoods in terms of ambiance (as being, e.g., "creative", "cozy"). Based on a resident survey, we found that ambiance scores matched respondents' expectations (§4).
- *Ambiance Prediction from social media images.* We used computer vision techniques to produce, for each image tagged with ambiance terms, a concept vector, which is a feature vector describing the image's content. We designed a regression model that predicted a neighborhood's ambiance based on the aggregation of the concept vectors of that neighborhood's images. Our regression model showed an accuracy as high as $R^2 = 0.80$ for some ambiance dimensions such as "stylish" or "quiet".
- *Distinctive Neighborhood Ambiance Discovery.* Finally, since different neighborhoods might express the same ambiance in different ways, we built a model that automatically identified the pictures most representative of "what an ambiance means in a specific neighborhood" (§6). With this method, we uncovered hidden corners of beauty, coziness, and joyfulness in overlooked parts of the city of London such as Stratford and Syon.

Our work has practical implications in a variety of contexts that go from advertising to urban planning (§7). The project's material is available at www.goodcitylife.org/ambiance.

2 RELATED WORK

Recent work has used social media to map urban characteristics at scale. Saeidi *et al.* used Twitter data and question and answering platforms to predict 62 demographic attributes for neighborhoods of London [28]. By matching taxonomically-organized terms with Flickr tags, other researchers have created sensory maps of smells and sounds for various cities across the world [1, 25]. Inspired by Milgram's psychological maps [17], past research has also used large-scale crowdsourcing platforms to collect and analyze judgments regarding inequality [29] and psychological prominence of urban areas [24].

Earlier urban computing studies used automatic image analysis tools to predict, for example, urban perceived safety [19] and changes in perceived safety [18]. Other work has focused on the vitality of cities [5] and neighborhoods [4], using both mobile phone traces and visual analysis. Computer vision techniques have also been used to understand what makes cities beautiful, quiet, and happy [7, 23], and to estimate ethnicity diversity through face analysis from social media data [31]. Few have used computer vision to predict some form of ambiance in urban environments. For example, Redi *et al.* [26] exploited the visual information of Foursquare user profile pictures to automatically predict the ambiance of the venues those users would tend to go to. Unlike previous work, we used social media images to predict ambiance of urban areas, and we did so by using accurate yet interpretable computer vision techniques based on deep learning.

Kafsi *et al.* [14] applied geographic hierarchy models to Flickr photos to discriminate between tags that characterize a neighborhood and those that characterizes surrounding regions. With a similar goal in mind, we built upon the work of Doersch *et al.* [6] and designed a model that is able to discover the most distinctive pictures of each neighborhood's ambiance. While the aim of Doersch *et al.*'s work was to find the most representative urban elements of a city, we instead aimed at capturing distinctive and multi-level ambiance traits of an urban area.

3 AMBIANCE TAXONOMY

We defined a 3-level taxonomy of urban ambiance: the first and more general level consists of **categories** of factors impacting ambiance (e.g., physiological, psychological); the second and intermediate level consists of **dimensions** (e.g., senses – "sound", perceptual characteristics – "joyful"); and the third and more specific level consists of **terms** characterizing each dimension (e.g., "loud", "cheerful"). Let us detail each level next.

Categories. Dupagne and Hegron [8] provided the most comprehensive definition of urban ambiance: *"The outcome of a complex composition of* ***physical****,* ***physiological****,* ***psychological****,* ***sociological*** *and* ***cultural*** *criteria"*. We populated the first level of the taxonomy with 5 *categories* mirroring the five aspects present in that definition.

Dimensions and terms. To associate each of the five categories with a list of *terms*, we identified terms that reflect ways in which people describe ambiances of urban spaces, and we did so with the help of three sources:

- *The Spirit of London.* A book [3] that describes London neighbourhoods through stories of people, buildings and places. While reading the entire book, we extracted all the 100+ adjectives (e.g., "bourgeois") that the author associated with features of the built environment such as streets and neighborhoods.
- *Indoor Ambiance Vocabulary.* We enriched the vocabulary using a set of 72 ambiance terms proposed by Graham and Gosling in a work in which they evaluated the ambiance of cafes and bars in Austin, Texas [9]. These terms described 3 different aspects of a place: the vibe (e.g., "gloomy"), the activities (e.g., "pick-up") and the personality of the patrons (e.g., "open-minded").
- *Airbnb Neighborhood Tags.* Airbnb provides a description for each neighborhood in twenty three cities[1]. Each neighborhood is described by a set of editorial tags and a list of user-generated words (e.g., "lively", "shopping", "West End"). From the full set of tags, we retained only words that are sufficiently general rather than being city-specific, and this filtering left us with 166 distinct words.

We then grouped the resulting terms into higher-level dimensions, and assigned each dimension to a category. We did so through Open Card sorting [32]. More specifically, we wrote down all the

[1] Airbnb Neighbourhoods: https://www.airbnb.co.uk/locations

Table 1. Neighborhood Ambiance Taxonomy. Ambiance is a multi-faceted concept and is characterized by physical, physiological, sociological, cultural, and psychological aspects. In this work, we focus on its *psychological* dimensions (which are thirteen) since reasonable proxies for these dimensions can be extracted from pictures.

	Dimension	Terms	
		Negative	**Positive**
Physical	*Land use (Greenness)*	Constructed, Built, Cemented	Green
	Location (Urbanization)	Isolated, Rural	Suburban, Urban, Central
	Size	Tiny, Small	Medium sized, Big, Panoramic
	Status	Under Construction, Developing, Re-Developing	Well-Established
Physiological	*Touch→ Temperature*	Cold	Warm
	Sight→ Light	Dark	Shiny, Bright
	Sight→ Colors	Colorless, Gray	Colorful
	Hearing → Sound	Silent, Quiet	Noisy, Loud
	Smell	Smelly	Scented
	Taste	Tasteless	Tasty
Sociological	*Cleanliness*	Dirty	Clean, Immaculate
	Gender	Masculine	Feminine
	Sexual Orientation	Straight	Gay
	Class	Working Class, Proletarian, Middle Class	Bourgeois, Professional, Upper Class, High Class, High-end
	Income	Poor, Cheap, Sustainable, Accessible, Gentrified	Wealthy, Pricey, Fancy, Rich, Luxurious, Expensive, Exclusive
	Etnicity	Monocultural, Homogeneous	Mixed, Diverse, Eclective, Ethnic, Exotic, Multicultural, International, Wordly
	Population Density	Deserted, Empty	Busy, Crowdy, Crowded, Dense, Over crowded
	Dwellers→ Age	Youthful	Retiree
	Dwellers→ Family Status	For-Singles	For-Couples, For-Families
	24-h use	Night, Diurnal, Daytime	24-hours, All-day
	7-d use	Weekend life, Weekday life	Full-7
	Public Uses (private, semi-public, public)	Residential, Industrial, Post-Industrial, Financial	Business, Commercial, Public
	Type of Activity (Night → Day)	Clubbing, nightlife, Party, Drinking, Dining	Leisure, Touristy, Sightseeing
Cultural	*Education*	Vulgar, Simple, Gritty	Smart,Intellectual, Sophisticated
	Secular	Orthodox, Puritan, Religious, Spiritual, Sacred	Profane, Secular
	Informality	Formal, Decorous, Conforming, Official	Frivolous, Casual, Spontaneous, Easygoing, Familiar, Free, Unofficial, Informal, Folksy
Psychological	*Creative*	Standardised, Typical, Dull, Banal, Ordinary, Normal, Regular, Conventional, Common, Undistinguished, Bland	Hipster, Artsy, Artistic, Bohemian, Special, Beatnik, Beat, Nonconformist, Unconventional, Hippie, Quaint, Unusual, Bizarre, Imaginative, Quirky, Strange, Creative, Eccentric, Crazy, Spacey, Spacy, Dizzy, Weird , Offbeat, Queer
	Unique	Homogenised, Monochrome, Clone-Street, Sterile, Noplace, Hackneyed, Corny, Trite, Predictable	Mysterious, Charming, Funky, Surprising, Baffling, Enigmatic, Confusing, Absorbing, Alluring, Engaging, Enticing, Captivating, unpredictable, Interesting, Unique, Distinctive, One of a Kind, Uncontaminated, Authentic, Different, Peculiar, Rare, Singular
	Beautiful	Nasty, Obnoxious, Distasteful, Awful, Grisly, Horrid, Ugly, unpleasant, Homely,	Attractive, Cute, Picturesque, Pleasant, Scenic, Lovable, Admirable, Nice, Pleasing, Pretty, Appealing, Lovely, Cool, Groovy, Idyllic, Beautiful, Gorgeous, Marvelous, Stunning, Wonderful, Superb, Amazing, Astonishing, Awesome, Brilliant, Fabulous, Excellent, Outstanding
	Joyful	Depressing, Doomed, Daunting, Wrecked, Condemned, Hopeless, Upsetting, Ghastly, Frightening, Ghostly, Terrible, Dreadful, Horrifying, Scary, Grim, Sad, Gloomy, Somber, Funereal	Cheery, Cheerful, Entertaining, Exciting, Effervescent, Merry, Happy, Joyful, Ecstatic, Joyous, Enjoyable, Elating
	Lively	Boring, Monotonous, Lifeless, Uneventful, Humdrum, Dreary, No-Hum	Lively, Dynamic, Social, Vibrant, Animated, Festive, Energetic, Active, Busy, Vital, Vigorous, Convivial
	Quiet	Hectic, frantic, Turbulent, Boisterous, Chaotic, Frenetic, Frenzied, heated, Tumultuous, Noisy, Frantic, Agitated, Hysterical	Calm, Chilled, Chilled out, Mellow, Serene, Peaceful, Placid, Relaxed, Relaxing, Peaceful, Tranquil, Breezy, Easygoing, Low Key
	Friendly	Unfriendly, Uncomfortable, Inhospitable, Unfavorable, Adverse, Hostile, Nasty, Bitter, Unsympathetic	Cozy, Comfortable, Friendly, Hyggelig, Comfy, Cushy, Comforting, Warm, Homey, Snug, Amicable, Welcoming, Receptive, Chummy
	Cosmopolitan	Local, Folk, Folkloristic, Rustic, Provincial, Pristine	Glocal, Cosmopolitan, Worldly, Global, International, Worldwide
	Popular	Unknown, Unexplored, Remote, Unrecognised, Secret, Hidden, Intimate, Secluded, Undiscovered, Private, Uncharted	Recognized, Accepted, Remembered, Acknowledged, Prestigious, Important, Popular, Illustrious, Notable, Famed, Iconic, Renowned, Famous, Well-known, Emblematic
	Modern	Ancient, Antique, Archaic, Timeworn, Venerable, Relic, Aged, Historical, Decrepit, Old, Victorian, Gothic, Colonial, Classical, Mature, Vintage, Dated, Passe'	Modernized, Modern, Young, Timeless, New, Contemporary, Fresh, Recent, Late
	Fancy	Subdued, Modest, Sober, Humble, Down-to-earth, Unobtrusive, Moderate, Unassuming, Timid	Extraverted, Outgoing, Showy, Grand, Kitsch, Magnificent, Palatial, Majestic, Ornate, Ostentatious, Pretentious, Sensational, Ambitious, Dignified, Glorious, Imposing, Opulent, Splendid, Monumental
	Stylish	Old-Fashioned, Nostalgic, Outdated, Retro, Antiquated, Behind the times, Demode, Dated, Plain, Basic, Simple	Well-Dressed, Trendy, Clean-cut, Hip, Glamorous, Stylish, Dressy, Polished Fashionable, Chic, Classy, Posh, Elegant, Swank, Exclusive
	Alternative	Puritan, Severe, Puritanical, Strict, Bigot, Austere, Rigid, Conservative, Traditional, Reactionary, Moderate, Regressive, Rightist	Liberal, Lefty, Imaginative, Flexible, Reformist, Tolerant, Progressive, Revolutionary, Alternative, Rebellious

terms in our taxonomy on sticky notes, and then asked two contributors to: (1) group the terms by synonyms and, more generally, by semantic similarity; (2) add to the so-grouped synonyms (e.g., "comfortable" and "friendly") their corresponding antonyms (e.g., "unfriendly" and "hostile"); and (3) assign a label to each of the synonym-antonym groups. That label had to be one of the ambiance *dimensions* (e.g., cozy) which, in turn, fell into a more general ambiance *category* (e.g., cozy belongs to the psychological category). Our contributors were allowed to assign a term to multiple dimensions, and they did so only at times (e.g., the term "noisy" was placed in the 'hearing →

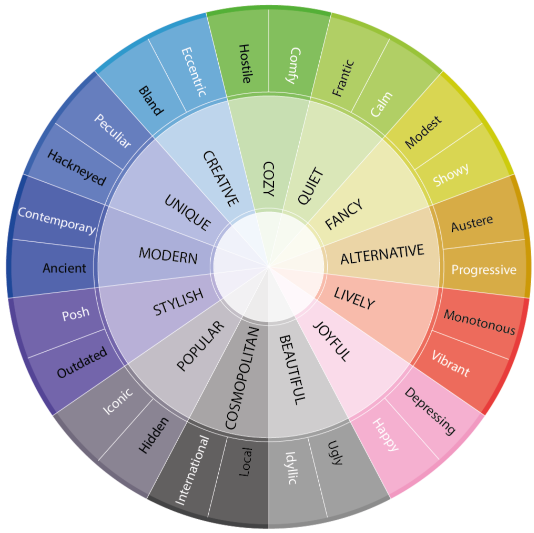


Fig. 1. Wheel representation of psychological ambiance taxonomy. The color of each ambiance dimension reflects the most dominant color of Flickr images tagged with terms related to that dimension.

sound' dimension of the physiological category, and in the 'sound' dimension of the psychological category).

As a result, we obtained, for each category, a set of ambiance dimensions, and for each dimension, a set of terms labeled according to their positive or negative associations with it (Table 1). For example, the physiological category (which captures the properties related to the five human senses) contains the terms "silent", "quiet", "noisy", and "loud". The first two were labeled as negative (i.e., negatively associated with the presence of sound), while the latter two as positive.

This work focused on what can be quantified from geo-referenced picture tags, and not all the categories in Table 1 can be quantified from tags. Indeed, the only one that can be satisfactorily quantified is the *psychological* category. That is because it contains terms that either: a) are generally mentioned when speaking about urban ambiance; or b) refer to concepts that can be easily extracted from pictures. However, a visual inspection of Table 1 offers a wider definition of urban ambiance, and this breath should facilitate future research on the unexplored categories, which require data other than geo-referenced pictures. Take the sociological category, for instance. It contains two dimensions called 'gender' and '24-h use'. These two dimensions might well be studied upon data from mobile phone records from which temporal mobility patterns and user-specific information can be extracted. Such studies might result in novel findings concerning gender and temporal uses of urban spaces.

To sum up, this paper focuses on the 13 ambiance dimensions that compose the "psychological" category: *Creative, Unique, Beautiful, Joyful, Lively, Quiet, Cozy, Cosmopolitan, Popular, Modern, Fancy, Stylish, Alternative*. To ease readability, we summarized those dimensions, as well as one example of a positive/negative term for each dimension, in the Ambiance Wheel in Figure 1. We also offer sample pictures for the positive aspect of each dimension and the negative one in Table 2. Next, we set out to ascertain whether our taxonomy had external validity and whether it matched the perceptions that locals had of their neighborhoods when applied to the analysis of London data.

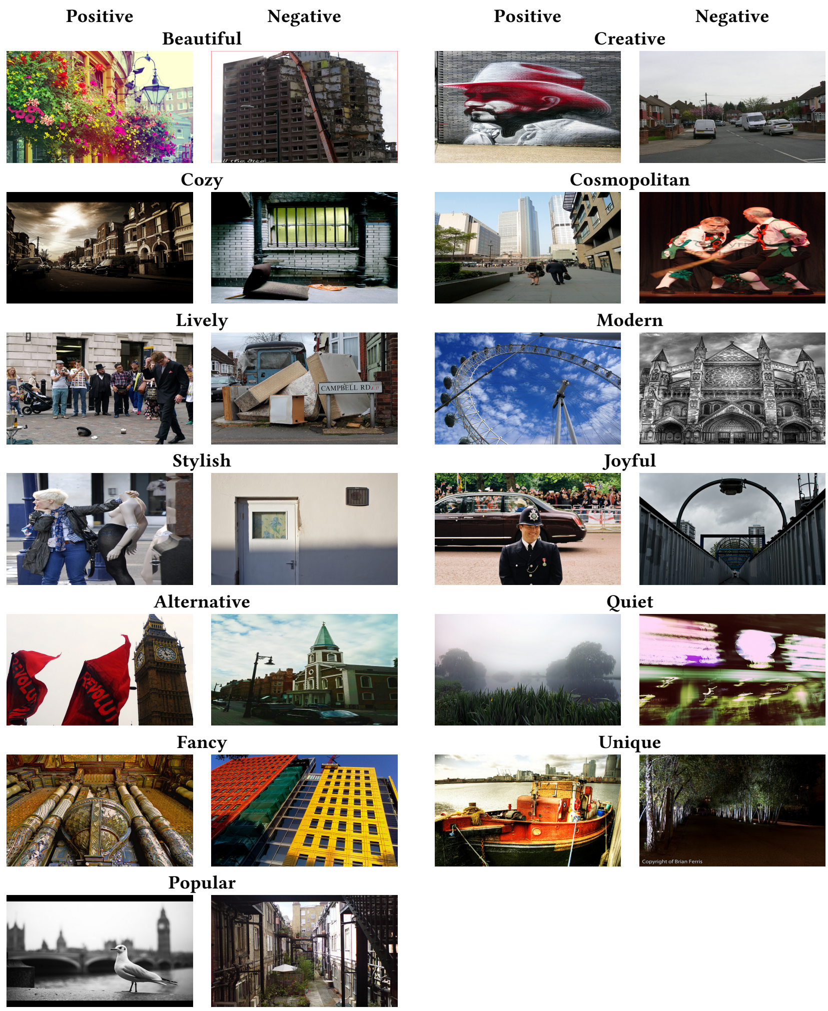


Table 2. Examples of images for positive and negative facets of each psychological ambiance dimension.

4 MAPPING AMBIANCE

To that end, we designed a methodology that uses social media data (§4.1) to assign ambiance scores to neighborhoods in London (§4.2), and evaluated its effectiveness with a user study and a qualitative analysis (§4.3).

4.1 Gathering Social Media Data

We collected social media data against which ambiance words were then matched. We focused on Flickr rather than on other social media platforms because Flickr content tends to be geographically salient: geo-located Flickr pictures are related to the places in which they have been taken more than, for example, what geo-located tweets would be [22]. Out of the set of all the public[2] geo-referenced Flickr pictures, we selected a random sample of about 3.5M public photos taken over

[2]Flickr API: www.flickr.com/api

	ward	borough
max number of tags per ward/borough	196,302 (St James)	307,268 (Westminister)
max number of images per ward/borough	130,427(St James)	209,806 (Westminister)
min number of tags per ward/borough	2 (St Michael)	1,832 (Sutton)
min number of images per ward/borough	2 (St Michael)	1,371 (Sutton)
average number of images per ward/borough	1,183	21,906
average number of tags per ward/borough	1,728	31,997
total number of images	722,912	
total number of tags	1,055,924	
average number of tags per image	1.5	

Table 3. Summary of our dataset.

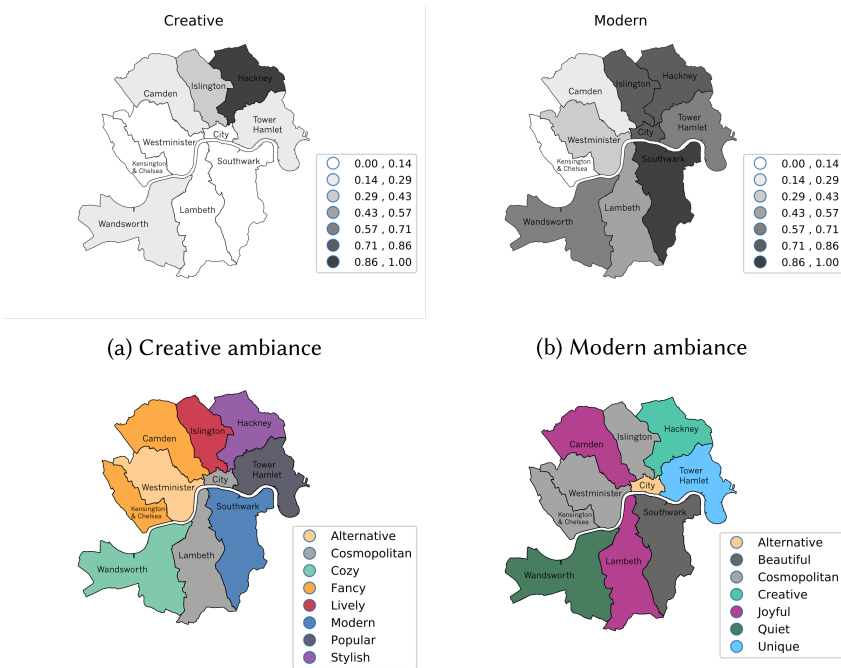
a 10-year period (2005-2015) within the bounding box of Greater London. For each picture, we collected the tags attached to it by its owner. There were 1.2M whose tags matched at least one term in the taxonomy. We made this data available to the community in agreement with the Flickr terms of service by publishing the set of photo identifiers in our collection[3]. To smooth out over-representation biases created by power users or large-scale events, we randomly sampled only γ pictures taken during a specific day, where γ was the average number of pictures taken per day during the whole period covered by our dataset (around 10 years). After this step, we were left with around 720K images, each with, on average, 1.5 tags. There were 32 boroughs (out of 32 London Boroughs) and 610 wards (out of 625 London Wards) covered by ambiance-related tags. As Table 3 details, the average number of images was 21,906 (per borough), and 1,183 (per ward), while the average number of tags was 31,997 (per borough), and 1,728 (per ward).

4.2 Scoring Neighborhood Ambiance

In this section, we present the methodology we used to calculate neighborhood ambiance scores. To briefly understand the methodology, consider that we took a neighborhood's pictures, e.g., those in Camden Town, took their user tags, i.e., the words used by Flickr users to describe these pictures (e.g., busy, social, vibrant, market, imaginative, people, noisy), and retained only those words matching the terms in our taxonomy (e.g., busy, social, vibrant, noisy, imaginative). Since the terms in our taxonomy are organized in positive/negative ones (e.g., 'busy, social, vibrant' are positive terms in the dimension 'lively', while 'noisy' is a negative term in the dimension 'quiet'), we counted, for all images in a neighborhood, the number of tags that were positive and those that were negative (e.g., for 'lively' in Camden Town, positive= 4570, negative= 136). We then obtained a positive/negative ambiance score by dividing this quantity by the total number of tags in a neighbourhood (e.g., for 'lively' in Camden Town, positive=$4570/11844 = 0.38$, negative=$136/11844 = 0.01$). We then normalized positive and negative ambiance scores by subtracting the mean positive/negative ambiance score over all neighbourhoods and dividing by the standard deviation (e.g., for 'lively' in Camden Town, positive=$\frac{0.38-0.1}{0.11} = 2.55$, negative=$\frac{0.01-0.04}{0.08} = -0.38$). We finally computed a neighborhood ambiance score as the difference between the normalized positive score and the normalized negative score (e.g., for Camden Town, lively=$2.55 - (-0.38) = 2.93$). This is just a sketch of the methodology. Let us detail it next.

Image Ambiance. We built a corpus of images annotated with positive/negative ambiance scores. Since for each of the 13 ambiance dimensions, we had 2 facets (positive and negative), we needed to separate the terms that were positively associated with ambiance j (T_j^+) from those that were negatively associated with it (T_j^-). As such, for each image i, we produced a *positive image ambiance*

[3]www.goodcitylife.org/ambiance



(a) Creative ambiance

(b) Modern ambiance

(c) The most prominent ambiance.

(d) The second most prominent ambiance.

Fig. 2. Ambiance maps of the ten central London boroughs. (a) Creative ambiance (darker shades reflect higher values); (b) Modern ambiance (darker shades reflect higher values); (c) most prominent ambinace; (d) second most prominent ambiance.

vector $A^+(i) = \{a_j^+(i)\}_{j=1}^{13}$ in which the j^{th} element corresponds to the number of i's tags that were positively associated with ambiance j (i.e., tags in T_j^+). We did the same for the negative and produced a *negative image ambiance vector* $A^-(i) = \{a_j^-(i)\}_{j=1}^{13}$.

Neighborhood Ambiance. Ambiances of geographic areas were then estimated by aggregating the ambiance vectors of images that are geo-located within that area. For the city of London, we aggregated pictures at two different spatial granularities: wards and boroughs. We mapped all the Flickr pictures in our dataset to the ward w and borough ω they fell into. There were 32 boroughs (out of 32) and 610 wards (out of 625) in which at least 1 image had an ambiance-related tag. For a ward w (or borough ω), we built positive and negative ward ambiance vectors by performing the element-wise addition of all the image ambiance vectors for that ward and by then normalizing the resulting scores by the total number of geo-referenced tags $T(w)$ in the ward. Formally:

$$A^+(w) = \{\alpha_j^+(w)\}_{j=1}^{13}; \qquad \alpha_j^+(w) = \sum_{i \in w} \frac{a_j^+(i)}{T(w)}. \tag{1}$$

To merge positive and negative ambiance facets into single score, we computed a normalized z-score vector $A(w) = \{\alpha_j(w)\}_{j=1}^{13}$, which balanced the two contributions:

$$\alpha_j(w) = \frac{\alpha_j^+(w) - \mu(\alpha_j^+)}{\sigma(\alpha_j^+)} - \frac{\alpha_j^-(w) - \mu(\alpha_j^-)}{\sigma(\alpha_j^-)}, \tag{2}$$

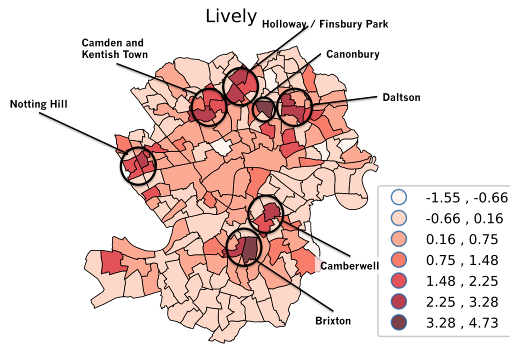


Fig. 3. The most and the least 'lively' wards in central London.

where $\sigma(\cdot)$ and $\mu(\cdot)$ correspond, respectively, to the standard deviation and the mean of the ambiance scores across wards. The same procedure was repeated at borough level.

The ward vectors capture the city ambiance as perceived by the sum of all users who have taken pictures within its boundaries, with no distinction between tourists and residents. We are interested in characterizing the typical spirit of an area without excluding any user category; future work can easily focus on specific groups using the same methodology.

Visualizing Neighborhood Ambiance. To inspect the spatial distribution of the different ambiance dimensions, we plotted the α_j scores on a set of choropleth maps. For the 10 most centric boroughs of London, we drew the maps of the two most complementary ambiance dimensions (creative and modern, Figure 2 (a-b)) and two maps of the first and second most prominent ambiances (Figure 2 (c-d)). Results met expectations. Hackney's most prominent ambiances were "creative" and "stylish": indeed Hackney hosts many fashion shops and art galleries. Kensington and Chelsea was labeled as "fancy" (it is indeed an upper-class area) and "cosmopolitan" (Notting Hill is part of it and has a large Caribbean community). The Borough of the City of London is the financial heart of the city and was labeled as "cosmopolitan" (first ambiance) and, less expectedly, "alternative" (second most prominent ambiance). This latter result could be explained through visual inspection: most of the pictures labeled as "alternative" depicted large-scale public gatherings such as protests. Finally, consider the "lively" ambiance dimension (Figure 3). The most lively areas turned out to be street markets and nightlife areas such as Camden Town, Brixton, Notting Hill, and Daltson (a recently gentrified area in East London).

4.3 Evaluating Neighborhood Ambiance Scores

To evaluate our ambiance scoring, we designed and administered a survey. To ensure quality responses, we recruited people who have lived in London for at least 5 years and selected the 8 most central and well-known London boroughs: Westminster, Camden, Hackney, Southwark, Lambeth, Kensington & Chelsea, City of London, and Islington. After reporting their age and gender, respondents were asked to give tips to an hypothetical friend who has to move to London and has to choose where to live based on specific needs. We selected these needs to reflect the four ambiance dimensions of "creative", "lively", "cosmopolitan", and "modern"[4]. Practically, we asked the following question: *"Which borough with a j vibe would you recommend to your friend?"*. Participants could select exactly one borough or, if they were not sure, they could select the option

[4]We selected these four ambiances as they are diverse between and are intuitive, therefore responders could easily rationalize their meaning in the context of a survey.

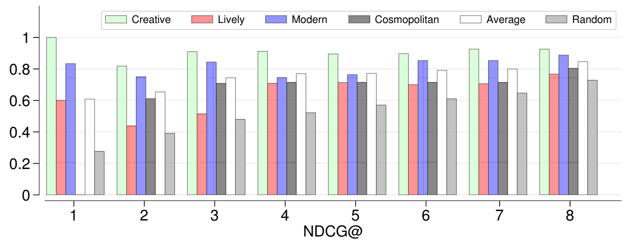


Fig. 4. The similarity (in terms of NDCG) between our borough ranking and the ground-truth ranking for five ambiances.

"none or any of them". We collected 27 responses (48% female, from 18 to 50 years old). For each borough-ambiance pair $\{\omega, j\}$, we computed a ground-truth score $\hat{\alpha}_j(\omega)$, which is the percentage of respondents that indicated borough ω as the preferred one for ambiance j.

For each ambiance j, we evaluated whether the ranking of boroughs generated by data-driven scores $rank(\alpha_j(\omega))$ is similar to the ranking of boroughs as chosen by the users $rank(\hat{\alpha}_j(\omega))$. To ascertain that, we computed the normalized discounted cumulative gain (NDCG), which is a measure that is specifically designed for non-binary notions of relevance, and is used to measure the similarity of two ranked lists in terms of how the top-k elements in the two lists are ranked in similar positions. If NDCG is 1, then the ranking is perfect: our data-driven ranking is the same as the ranking produced by our respondents.

In Figure 4, we report the results in terms of NDCG@k with $k = 1, \ldots, 7$. We compared the results with a random rank baseline. The baseline performed very poorly in terms of NDCG, and, as expected, its performance increased only with k. Unlike the baseline, our ambiance scoring method replicated very well our respondents' preferences: for some ambiances, where tags were unique and distinctive, such as for *modern* (NDCG(1)=0.8) and *creative* (NDCG(1)=1), our ranking almost perfectly matched the survey respondents' ranking. However, for other ambiances characterized by more generic tags (e.g., for *lively*), the ranking was less than ideal but always well above the baseline.

5 DESCRIBING AND PREDICTING AMBIANCE FROM PICTURES

After being able to map ambiance and ascertain the effectiveness of such mapping, we were then ready to use computer vision to focus on the pictures. That is, to explore the relationship between an area's ambiance and the visual content of the area's pictures.

5.1 Visual Concept Scoring

Pictures from social media exhibit extremely varied visual properties due to factors such as lighting conditions, filters, and photographic techniques. This strong heterogeneity makes it difficult to train computer vision classifiers to identify ambiance from visual features without incurring in visual biases. For this reason, we decided to use deep learning to reliably extract visual concepts from the pictures and to then use these concepts to predict the pictures' ambiance dimensions. To that end, we ran, on each image, a deep-learning visual concept detector, which is similar to Krizhevsky et al.'s [15]. The detector was able to recognize up to 1750 distinct visual concepts (including scenes, objects and activities), which are organized in a 5-level taxonomy (the categorization of, for example, clownfish unfolds from 0 animal → 1 marine life → 2 fish → 3 anemone fish → 4 clownfish). To reduce sparsity, we considered levels 0 and 1 only, and mapped all the concepts at lower levels to

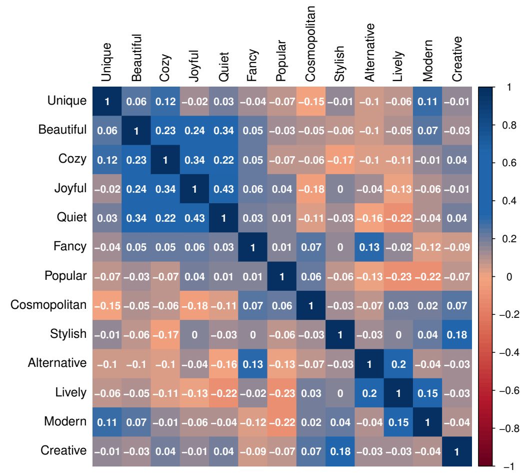


Fig. 5. Cross rank correlations between ambiances.

their ancestors on level 1. This reduced the number of concepts to 671. The detector was able to recognize visual concepts in the image and to output, for each visual concept v, a confidence score c_v in the interval $[0, 1]$. This score reflected the probability that the image actually showed that concept. We discarded all concepts with confidence < 0.5.

We described the visual content of an image i with a 671-dimensional *visual concept vector* $V(i) = \{v_l(i)\}_{l=1}^{671}$ where each element corresponded to the average confidence score c_l for a concept l. Similar to Eq. (1), image vectors could be aggregated at ward level to estimate the visual vector $V(w) = \{v_l(w)\}_{l=1}^{671}$ of a ward, whose elements are the sum of the confidence scores associated with each concept l normalized by the total number of images in the ward $I(w)$:

$$v_l(w) = \sum_{i \in w} \frac{v_l(i)}{I(w)}. \tag{3}$$

5.2 Explaining Neighborhood Ambiance

With those metrics at hand, we were then able to answer questions concerning ambiance and visual concepts. What makes a ward cozy, creative or lively? Which visual concepts are associated with which ambiance categories?

Ambiance-Ambiance Correlations. First, we looked at how ambiance dimensions were related to each other by computing the Spearman Rank correlation between pairs of ambiances across the 610 wards (Figure 5). We found positive correlations between “quiet”, “friendly”, “joyful”, and “beautiful” ambiances, all related to pleasant feelings. By contrast, we found fairly negative correlations between the “popular” ambiance and the “quiet” or “friendly” ambiances. Similarly, the “lively” ambiance was negatively correlated with “quiet”.

Ambiance-Concept Correlations. Next, we computed the rank correlation between ambiances and visual concepts. We report the results using word clouds, which show the top 50 visual concepts for each ambiance category (Table 4). As expected from the literature [23], pictures from “beautiful”, “joyful”, and “quiet” wards tended to show outdoor images and green areas with no cars. In “creative”

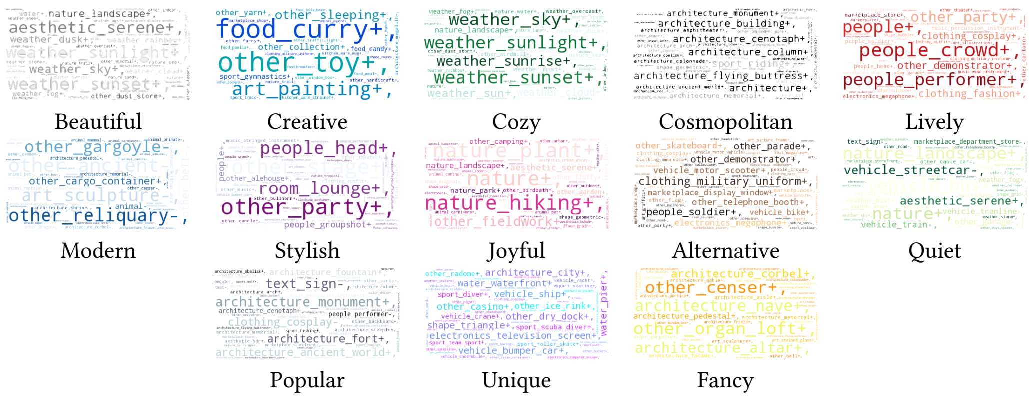


Table 4. Wordclouds showing the words that are associated with the highest correlations between visual concepts (e.g., aesthetics, people's head) and each ambiance dimension (e.g., beautiful, creative, cozy). The font size reflects correlation strength, and '+'/'-' signs reflect whether the correlations are positive or negative.

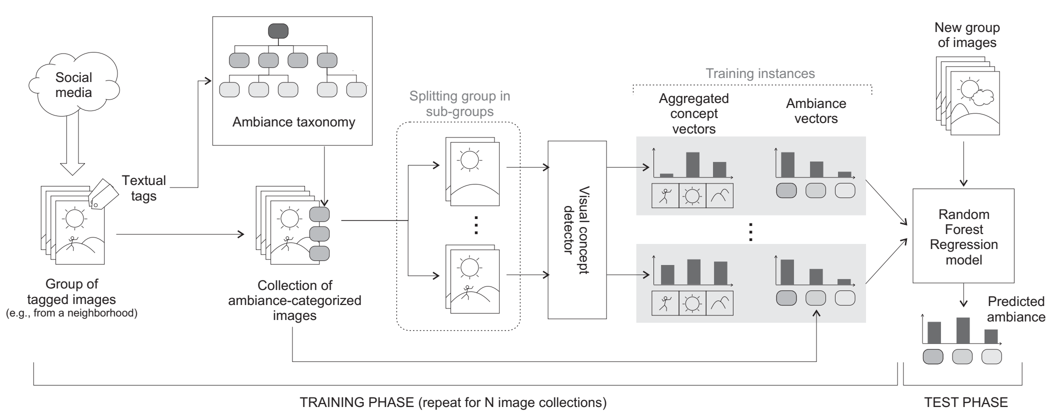


Fig. 6. Ambiance Prediction Task with Training Data Augmentation (left to right). Geo-referenced pictures are collected from social media (e.g., Flickr), and their tags matched with the terms in our ambiance taxonomy. The task is to predict a neighborhood's *ambiance* from the *concepts* extracted from the neighborhood's pictures. Since neighborhoods are limited (they are 610), then the resulting training set would be limited: it would only have 610 (concepts, ambiance) pairs. Therefore we need a step to augment the training set. To this end, we split a neighborhood's pictures into sub-groups (artificially augmenting the training points for that neighborhood), apply a visual concept detector to extract concepts from each sub-group's pictures, and apply a random forest regression classifier to (concepts, ambiance) pairs at sub-group level.

areas, one could find food markets and art galleries, while pictures of "stylish" neighborhoods showed people gathering or partying. Images of "cozy" wards showed outdoor landscapes at sunset and sunrise. "Cosmopolitan" areas such as the City of London were all about architectural elements. "Lively" wards were generally associated with the presence of people and street artists, while "alternative" ones were associated with demonstrations and protests. While "modern" areas were associated with the absence of classical architectural elements, "popular" ones were associated with touristic attractions, and "fancy" ones with palatial architecture. Finally, the "unique" ambiance dimension showed low correlations as there was no specific visual pattern associated with it.

5.3 Predicting Ambiance

To understand the extent to which the local production of visual content was predictive of a neighborhood's ambiance, we designed a machine learning framework that predicted each ambiance dimension from our concept vectors (Fig. 6).

In terms of ambiance, each ward was characterized by a set of 13 ambiance scores, captured in the ward's ambiance vector. As a result of Eq. (2), these scores were continuous. If we binned them into discrete categories (e.g., high, medium, low), we would have lost the subtle ambiance differences among wards. As such, predicting raw and continuous ambiance scores with a regression framework seemed to be the best solution.

The restricted number of data points (610 wards) called for a data augmentation step (fully explained in Fig. 6). Similar to previous work [26], we performed the augmentation by randomly partitioning the set of images $I(w)$ in each ward w into g_w equally-sized sub-groups $I_1(w) \dots I_{g_w}(w)$, and by re-computing the concept vectors for these sub-groups. We reduced the possibility of grouping near-duplicates by selecting non-overlapping image groups. We labeled each sub-group concept vector for each ambiance j with the ward-level ambiance score $\alpha_j(w)$. To avoid geographic biases (e.g., the risk of associating a certain ambiance with a highly photographed location), we made sure that our group concept vectors were defined to be sufficiently general. These vectors contained descriptions of the objects and patterns that were sufficiently general to be location independent - for example, they described the general concept of a 'bridge' rather than the location-specific concept of 'London Bridge'.

We then wanted to predict ambiance from concept vectors. To that end, we trained a regression classifier able to output a predicted ambiance score $\alpha'_j(g, w)$ for the ward where the group g had been drawn from. To perform regression on groups of images, we trained a Random Forest Regressor with 30 trees. We chose Random Forest because it has been shown to be the most effective supervised learning algorithm for high-dimensional problems [2]. We divided all our 722K images into five partitions, used three partitions for training, one for validation, and one for test. Before doing any training, we needed to set the best parameters (mainly the model's number of trees), and we did so by performing grid-search on the training and validation sets. We accordingly set the parameter values, trained the model on validation+training sets, and tested on the remaining partition. We did this iteratively for all the five partitions, while keeping track of the R^2 coefficient, which reflected how well the regression line fitted the observed data points. We reported the results in Figure 7 as average R^2 on the five test sets with corresponding error bars.

In Figure 7, R^2 varies with group size. We can observe that having one image ($x = 1$) was not predictive of ward ambiance, while it got better when considering multiple images. However, the accuracy increased only up to a point (x=100). For larger sets of images, the accuracy got worse, and the prediction unreliable (which was reflected in the wide error bars). That might have been because the greater the number of images, the higher the noise introduced in the average concept vector, and the smaller the number of data points used for prediction. Some ambiances such as "quiet" were easy to predict because the corresponding pictures had quite distinctive visual cues. By contrast, others such as "cosmopolitan" are more multi-faceted and, as such, are harder to predict.

6 WHAT MAKES SHOREDITCH CREATIVE?

Not all neighborhoods with the same ambiance have the same visual appearance: what makes Shoreditch "creative" might be different from what makes Clerkenwell so. That is why we built a model that was able to determine which of a neighborhood's images were most strongly associated with an ambiance dimension, and that allowed us to discovered how the same ambiance might have been differently paraphrased across neighborhoods.

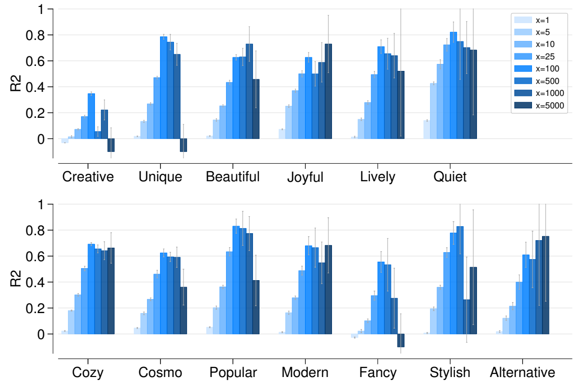


Fig. 7. Ambiance Prediction Task: average R^2 for ambiance classification of image groups, per ambiance and per group size. Error bars represent the maximum and the minimum deviations from the average derived from a 5-fold cross-validation.

6.1 Discovering Ambiance-Specific Neighborhood Pictures

More specifically, for each ambiance j, we first designed a multi-class classifier that, given an image geo-located in one of the top-10 wards for that ambiance j, was able to automatically predict its ward location. We were then able to determine the most representative images of ambiance j at each ward, and we did so by relying on the classifier's learned feature space.

Ward Classifier. For each ambiance j, we selected the 10 wards with the highest ward-ambiance scores $\alpha_j(w)$. We then described each image i with its visual concept vector $V(i)$. With a setup similar to the one described in §5.3, we then trained a Random Forest classifier that, given a picture i, returned a predicted ward $w'_j(i)$ and the probability $P(w_j|i)$ of picture i being in w_j. We split images into 5 partitions, iteratively use 3 partitions for training, 1 for validation, and 1 for test. By computing the average accuracy on the test set, we found that a random classifier had an accuracy of 10%, while our approach achieved 37%, with a 0.07 standard deviation.

Self-learning training. Similar to previous work [6], to select representative examples that both (1) frequently occur in a given ward and (2) are visually distinctive for that ward, we followed a self-learning approach that unfolded in three steps.

Data Cleaning. With the model learned in the initial training set, we computed, for an image, its distinctiveness for its ward, which is $P(w_j|i)$. The higher the probability of an image i to belong to a ward w_j, the more likely it is that i represents the unique pattern that makes w_j distinguishable from others. For each ward w_j, we then filtered out the images that are not highly distinctive, that is, for which $P(w_j|i) < 0.8$.

Classifier Re-Training. With this training set, we retrained the ward classifier. We iteratively performed such classifier-retraining procedure for 3 iterations as in previous work [6]. The resulting classifier learned to identify each ward's most distinctive images from the test set.

Discovery Phase. We selected, for each ambiance, the pair of wards with lowest mutual misclassification rate in the training set. This ensured that, for each ambiance, we selected those wards that were visually distinctive to one another. We classified and ranked images in the the test set held-out at each partition according to their probability of belonging to ward w_j. Indeed, the higher $P(w_j|i)$, the higher the distinctiveness of image i for ambiance j in ward w.

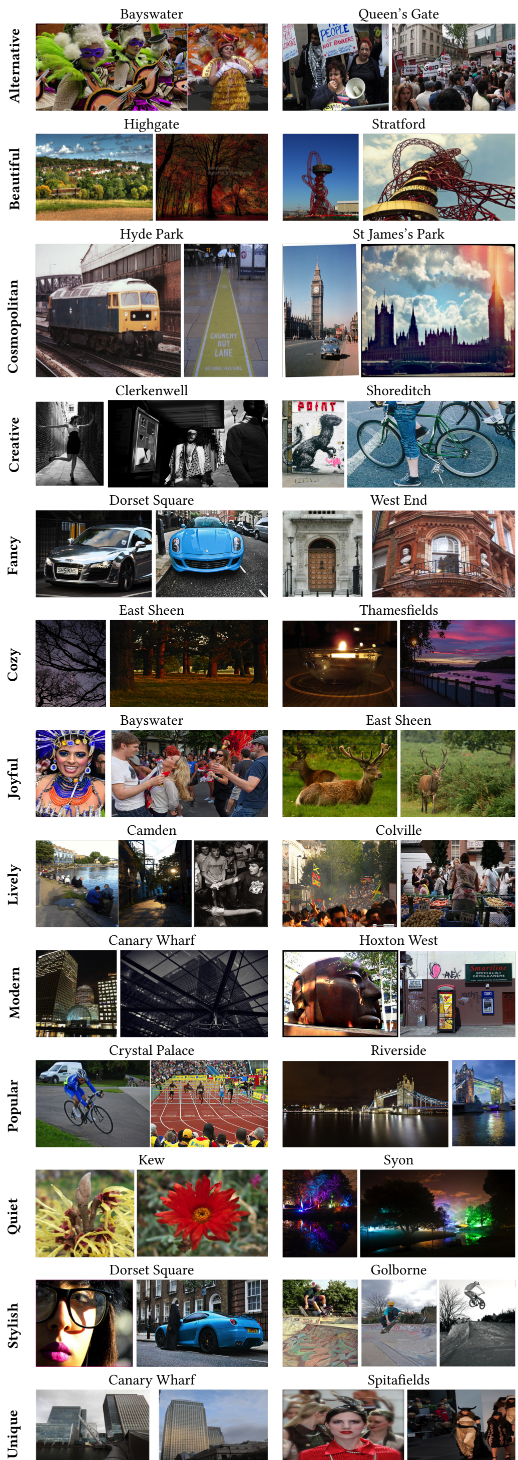


Table 5. The most distinctive images for each ambiance.

6.2 Same Ambiance, Different Neighborhoods

We contrast pairs of wards by determining their two most distinctive images for each ambiance j (Table 5). Each pair of wards represented the two most visually diverse wards for a given ambiance.

Alternative. Pictures of alternative wards showed large-scale gatherings. These were either protests (like in Queen's Gate) or celebrations such as the Notting Hill Carnival (in Bayswater).

Beautiful. Beauty in London took many different forms, from the traditional parks of Highgate to hidden gems such as the Anish Kapoor's Orbit Tower in Stratford.

Cosmopolitan. The top cosmopolitan wards tended to have similar visual attributes. The corresponding pictures generally depicted public gathering areas such as stations (e.g., Victoria Station near Hyde Park) and touristy places (e.g., St James's Park).

Creative. Creativity is in the DNA of East London, and pictures reflected that. Clerkenwell was associated with fashion and design, and Shoreditch with graffiti, street artworks, and the "hipster lifestyle" (bikes, beards, and cafes were prominent in the pictures).

Fancy. Fancy neighborhoods were found in the Western areas and were associated with a wide variety of elements, from the palatial architecture of the West End to high-end cars in Dorset Square.

Cozy. Coziness was generally associated with warm colors and sunsets in parks (e.g., East Sheen), or quiet, relaxing landscapes (e.g., Thamesfield).

Joy. The Joyful ward classifier was the most accurate and associated joyful ambiance with calm and relaxing places (e.g., East Sheen), and with large-scale gatherings (e.g., the Notting Hill Carnival).

Liveliness. Camden's unique liveliness was associated with its canal boat life, the crowds of the diurnal Camden Market, and those of the Camden-style nightlife. Colville, one of the most lively area in the heart of Notting Hill, was associated with its annual carnival celebrations.

Modern. Modernity was associated with different parts in the city. For example, it was associated with Hoxton which is a gentrified area in North-East London, and with Canary Wharf which is the financial area of the city.

Popular. Popular areas were associated not only with well-known touristic landmarks (e.g., the riverside) but also with large-scale events for locals (e.g., sport events at Crystal Palace).

Quiet. Quiet places included Kew Gardens, with its unique flora species, and Syon Park, with its 1-month light display called "Enchanted Woodland".

Stylish. Style was associated with a variety of concepts, from 'posh' in Dorset Square to 'free-style' in the Skate Park at Goldborne.

Unique. The most unique neighborhoods were those with unique architecture (e.g., Canary Wharf) and those with unique events (e.g., Spitafiled Alternative Fashion Show).

7 DISCUSSION AND CONCLUSION

We designed an end-to-end system that captured and visually characterized neighborhood ambiance at an unprecedented scale. This work has resulted in theoretical and practical implications.

Theoretical Implications. We contributed to the literature of urban planning by designing the first complete taxonomy of urban ambiance. The taxonomy is literature-driven and carefully curated. Our work explored the psychological category, while leaving the remaining categories in Table 1 for future work: social media researchers and urban planners could use them to relate ambiance to socio-economic indicators, social psychologists to study urban perceptions at scale, and computer vision researchers to produce high-quality training data for urban ambiance.

Practical Applications. In addition to characterize ambiance of places, our method can be applied to user profiles (say, a person's Instagram pictures) [26]. This would make it possible to recommend places based on one's personal preferences for ambiance. Advertising campaigns and marketing initiatives could make use of their potential clients' favorite ambiances, and location-based applications such as Facebook places, Google maps, Yelp and AirBnb could do just the same.

Limitations. We modeled the "average" ambiance perception at neighborhood level. However, ambiance perception can change from subject to subject and from country to country [13]. Data from social media [27, 33], including photo sharing platforms [34], carry a number of biases that make it difficult to accurately predict the ambiance of an individual picture without considering the socio-cultural factors of the picture creator and his/her context. Those biases are mitigated by aggregating pictures and analyzing them all together. Indeed, by aggregating crowd-sourced geo-spatial data, previous work has captured subjective and intangible properties of the urban space [12, 25], and our work borrowed from that stream of research. In the future, we envision a complementary study about the relationship between people demographics and ambiance perception. Another limitation comes from the fact that our ambiance predictor is composed of several steps: (1) concept feature extraction; (2) feature aggregation; and (3) random forest regression. Although very effective, each of these steps carries its own biases: (1) object detectors come with an inherent prediction error; (2) feature aggregation might destroy important image patterns; and (3) random forest models might incur in data over-fitting. To overcome these issues, we plan to build an end-to-end deep learning architecture that takes as input the set of a neighborhood's pictures and gives as output the neighborhood's ambiance score. Lastly, our taxonomy is not final – it is just a starting point upon which researchers can add new dimensions or enrich existing ones. That could well be done using the very same methodology this paper has proposed.

REFERENCES

[1] Luca Maria Aiello, Rossano Schifanella, Daniele Quercia, and Francesco Aletta. 2016. Chatty Maps: Constructing Sound Maps of Urban Areas from Social Media Data. *Royal Society open science* 3, 3 (2016), 150690.

[2] Rich Caruana, Nikos Karampatziakis, and Ainur Yessenalina. 2008. An Empirical Evaluation of Supervised Learning in High Dimensions. In *International Conference on Machine Learning (ICML)*. ACM, 96–103.

[3] P. Cohen-Portheim and S. Jenkins. 2011. *The Spirit of London*. Pavilion Books.

[4] Marco De Nadai, Jacopo Staiano, Roberto Larcher, Nicu Sebe, Daniele Quercia, and Bruno Lepri. 2016. The Death and Life of Great Italian Cities: A Mobile Phone Data Perspective. In *International Conference on World Wide Web (WWW)*. ACM, 413–423.

[5] Marco De Nadai, Radu Laurentiu Vieriu, Gloria Zen, Stefan Dragicevic, Nikhil Naik, Michele Caraviello, Cesar Augusto Hidalgo, Nicu Sebe, and Bruno Lepri. 2016. Are Safer Looking Neighborhoods More Lively?: A Multimodal Investigation into Urban Life. In *International Conference on Multimedia (MM)*. ACM, 1127–1135.

[6] Carl Doersch, Saurabh Singh, Abhinav Gupta, Josef Sivic, and Alexei A. Efros. 2015. What Makes Paris Look Like Paris? *Commun. ACM* 58, 12 (2015), 103–110.

[7] Abhimanyu Dubey, Nikhil Naik, Devi Parikh, Ramesh Raskar, and César A Hidalgo. 2016. Deep learning the city: Quantifying urban perception at a global scale. In *European Conference on Computer Vision*. Springer, 196–212.

[8] Alex Dupagne and Gérard Hégron. 2002. Introduction. In *Workshop on Architectural and Urban Ambient Environment*.

[9] Lindsay T Graham and Samuel D Gosling. 2011. Can the Ambiance of a Place be Determined by the User Profiles of the People who Visit it?. In *International Conference on Web and Social Media (ICWSM)*. AAAI.

[10] William Harispuru. 2012. Positive Distractions in Urban Ambiance - Exploring the Elements that Help Synthesize Our Healthy Modern Experiences. In *Ambiances in action: International Congress on Ambiances*. International Ambiances Network, 473–478.

[11] Tony Hiss. 1990. *The Experience of Place: A New Way of Looking at and Dealing with our Radically Changing Cities and Countryside*. Vintage Books.

[12] Livia Hollenstein and Ross Purves. 2010. Exploring place through user-generated content: Using Flickr tags to describe city cores. *Journal of Spatial Information Science* 2010, 1 (2010), 21–48.

[13] Brendan Jou, Tao Chen, Nikolaos Pappas, Miriam Redi, Mercan Topkara, and Shih-Fu Chang. 2015. Visual Affect Around the World: A Large-scale Multilingual Visual Sentiment Ontology. In *International Conference on Multimedia*

(MM). ACM, 159–168.

[14] Mohamed Kafsi, Henriette Cramer, Bart Thomee, and David A. Shamma. 2015. Describing and Understanding Neighborhood Characteristics Through Online Social Media. In *International Conference on World Wide Web (WWW)*. ACM, 549–559.

[15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of ACM* 60, 6 (2017), 84–90.

[16] Stanley Milgram. 1970. The Experience of Living in Cities. *Science* (1970).

[17] Stanley Milgram. 1976. Psychological Maps of Paris. *Environmental psychology* (1976).

[18] Nikhil Naik, Scott Duke Kominers, Ramesh Raskar, Edward L Glaeser, and César A Hidalgo. 2017. Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences* 114, 29 (2017), 7571–7576.

[19] Nikhil Naik, Jade Philipoom, Ramesh Raskar, and César Hidalgo. 2014. Streetscore: Predicting the Perceived Safety of One Million Streetscapes. In *Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 779–785.

[20] Christian Norberg-Schulz. 1980. *Genius loci: Towards a phenomenology of architecture*. Rizzoli.

[21] Juhani Pallasmaa. 2014. Space, Place and Atmosphere. Emotion and Peripherical Perception in Architectural Experience. *Lebenswelt. Aesthetics and philosophy of experience.* 4 (2014).

[22] Daniele Quercia. 2013. Psychological Maps 2.0: A web engagement enterprise starting in London. In *International Conference on World Wide Web (WWW)*. ACM, 591–592.

[23] Daniele Quercia, Neil Keith O'Hare, and Henriette Cramer. 2014. Aesthetic Capital: What Makes London Look Beautiful, Quiet, and Happy?. In *Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*. ACM, 945–955.

[24] Daniele Quercia, Joao Paulo Pesce, Virgilio Almeida, and Jon Crowcroft. 2013. Psychological Maps 2.0: A Web Engagement Enterprise Starting in London. In *International Conference on World Wide Web (WWW)*. ACM, 1065–1076.

[25] Daniele Quercia, Rossano Schifanella, Luca Maria Aiello, and Kate McLean. 2015. Smelly Maps: the Digital Life of Urban Smellscapes. In *International Conference on Web and Social Media (ICWSM)*. AAAI.

[26] Miriam Redi, Daniele Quercia, Lindsay T Graham, and Samuel D Gosling. 2015. Like Partying? Your Face Says It All. Predicting the Ambiance of Places with Profile Pictures. In *International Conference on Web and Social Media (ICWSM)*. AAAI.

[27] Derek Ruths and Jürgen Pfeffer. 2014. Social media for large studies of behavior. *Science* 346, 6213 (2014), 1063–1064.

[28] Marzieh Saeidi, Alessandro Venerandi, Licia Capra, and Sebastian Riedel. 2017. Community Question Answering Platforms vs. Twitter for Predicting Characteristics of Urban Neighbourhoods. *arXiv:1701.04653* (2017).

[29] Philip Salesses, Katja Schechtner, and César A Hidalgo. 2013. The collaborative image of the city: mapping the inequality of urban perception. *PloS one* 8, 7 (2013), e68400.

[30] Myriam Servieres, Gwendoline LHer, and Daniel Siret. 2017. Mobile Devices and Urban Ambiances: How Connected Wearable Tools Change the Ways We Perceive and Design Public Spaces. In *Urban Design and Representation*. Springer, 207–216.

[31] Vivek Kumar Singh, Saket Hegde, and Akanksha Atrey. 2017. Towards Measuring Fine-Grained Diversity Using Social Media Photographs.. In *International Conference on Web and Social Media (ICWSM)*. AAAI.

[32] Donna Spencer. 2009. *Card Sorting: Designing Usable Categories*. Rosenfeld Media.

[33] Zeynep Tufekci. 2014. Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. In *International Conference on Web and Social Media (ICWSM)*. AAAI.

[34] Haipeng Zhang, Mohammed Korayem, David J. Crandall, and Gretchen LeBuhn. 2012. Mining Photo-sharing Websites to Study Ecological Phenomena. In *International Conference on World Wide Web (WWW)*. ACM, 749–758.

Received April 2018; revised July 2018; accepted September 2018