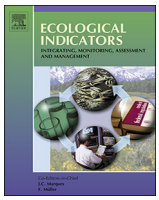
Ecological Indicators 108 (2020) 105722



Contents lists available at ScienceDirect

Ecological Indicators

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



Measuring urban landscapes for urban function classification using spatial metrics



Hanfa Xing[a,b,c], Yuan Meng[a,b,*]

[a] College of Geography and Environment, Shandong Normal University, Jinan, Shandong 250300, China
[b] Research and Development Centre for Big Geo-Data, Shandong Normal University, Jinan, Shandong 250300, China
[c] School of Geography, South China Normal University, Guangzhou, Guangdong 510631, China

ARTICLE INFO

Keywords:
Urban function
Classification
Urban landscape
Spatial metric
Urban form

ABSTRACT

The rapid extension of urban landscapes has substantially contributed to the widespread use of methodologies for the classification of urban functions. While previous research has focused solely on measuring the spatial arrangement of restricted urban landscape elements using spatial metrics, to date, the detailed landscape characteristics that enable different functions to be distinguished have not been discussed. This paper presents a novel framework for urban function classification by measuring urban landscapes using spatial metrics. To capture features related to urban functions rather than measuring limited physical configurations, spatial metrics that quantify multiple urban landscape elements and their interactions are proposed. Then, these metrics are evaluated based on their sufficiency and effectiveness in classifying urban functions using a partial analysis and analysis of variance. Finally, a conditional inference random forest approach is proposed to build an automatic urban function classification model with spatial metrics. The model was implemented in Futian District, Shenzhen, Guangdong Province. Using 8 spatial metrics based on 6 types of urban landscape elements, the results show a satisfactory testing accuracy of 0.757 and an overall classification accuracy of 0.813. The proposed method provides a novel framework for effectively delineating urban landscape characteristics and accurately classifying urban functions.

1. Introduction

Urban function refers to the purpose of areas of residential, commercial, and public uses. Urban function contains the physical configuration of a city that arises from changes in urban forms (Manonmani and Suganya, 2010) and is considered an important component in urban planning and resource management (Anderson, 1976; Matsuoka and Kaplan, 2008). Thus, the recognition of urban forms is essential for understanding and classifying the wide range and distribution of urban functions. In general, urban forms can be described as the spatial patterns of urban landscapes that consist of various fundamental elements (Abrantes et al., 2017). The conventional methods used to measure urban forms heavily rely on the analysis and simulation of statistical data, such as the population, commuting patterns and changes in employment (Bertaud and Renaud, 1997; Tsai, 2005). Such approaches are restricted to the acquisition of statistical data, leading to short time intervals of data related to urban characteristics. The proliferation of GIS-based data and remotely sensed images has provided more effective approaches for delineating urban forms (Liu et al., 2018; Weng, 2012; Zheng et al., 2014). However, the utilisation of these data sources has failed to capture the landscapes of urban forms that enable the detailed morphologies of urban functions to be distinguished.

In fact, urban landscapes inspired by landscape ecology are spatially organised into irregular landscape elements, such as buildings, green land, and roads. The morphology of these elements (e.g., shape, size, edge, and density) and their interactions (ratios between different elements) effectively provide quantitative information regarding urban forms (Barr and Barnsley, 1997). Existing research has attempted to investigate the relationship between urban landscapes and functions. Barnsley and Barr (1997) proposed a graph-based structural pattern recognition system that could infer different categories of urban functions from fine spatial resolution. From the perspective of temporal changes in urban landscapes and functions, Alberti and Marzluff (2004) analysed the complex interactions between urban patterns and ecosystems by developing a conceptual model to assess the resilience of urbanising regions. In addition, Baus et al. (2014) statistically modelled landscape pattern changes to gain a better understanding of urban dynamics. From the perspective of spatial changes in urban landscapes

* Corresponding author at: College of Geography and Environment, Shandong Normal University, Jinan, Shandong 250300, China.
E-mail addresses: xinghanfa@sdnu.edu.cn (H. Xing), mengyuan@stu.sdnu.edu.cn (Y. Meng).

https://doi.org/10.1016/j.ecolind.2019.105722
Received 13 August 2018; Received in revised form 12 July 2019; Accepted 9 September 2019
1470-160X/ © 2019 Elsevier Ltd. All rights reserved.

and functions, Kane et al. (2014) measured urban landscapes to analyse urban sprawl based on urban areas, fragmentation, shape complexity and diversity. Li et al. (2017) utilised a geographically weighted regression (GWR) to identify spatial heterogeneity in landscape patterns and the relationship between landscape patterns and urbanisation.

Numerous studies have been conducted to understand and quantify urban landscapes based on the strong relationships between landscapes and functions. In particular, spatial metrics, which were originally developed to describe the spatial components in ecological landscapes, are often utilised in these studies (McGarigal et al., 2002). Liu and Yang (2015) utilised spatial metrics originally designed to measure the size, shape and spatial arrangement of natural landscapes to characterise urban patterns. On this basis, the traditional spatial metrics have been further modified to quantify urban landscapes, including analysing morphological properties based on building-level blocks (Chen et al., 2014; Liu et al., 2017a) and utilising road-based metrics to define distinct urban topologies (Hermosilla et al., 2014). In addition, for large-scale urban landscapes, different levels of spatial metrics, such as the dispersion index and built-up density index, have been proposed for cities (Taubenböck et al., 2014). Despite the existence of these traditional and new spatial metrics, there are still no metrics specifically designed for inferring urban functions. Thus, directly applying the existing metrics may overestimate or ignore typical landscape characteristics that enable different functions to be distinguished.

To fill the gap between urban landscape quantification and functional classification, numerous approaches have been proposed to promote effective spatial metrics and automatically classify urban functions. With regard to delineating urban function regions, previous studies have utilised predefined regular pixels as basic units and applied spatial metrics to measure the landscape features in each unit (Montanges et al., 2015; Tu et al., 2018). However, these methods are restricted to regular units and, thus, fail to capture actual landscape patterns among different units. To address this issue, many approaches aim to quantify urban landscapes and delineate functions based on the spatial distribution of urban landscape elements. For instance, main roads from urban landscape elements have been considered to divide regions as basic units for spatial metrics measurements (Van de Voorde et al., 2011). On this basis, urban landscape element-driven spatial metrics are considered to depict the characteristics of urban functions from different aspects. Existing studies have proposed building-level spatial metrics, such as building adjacency, mean building size and building size standard deviation (Vanderhaegen and Canters, 2017; Xing and Meng, 2018). Although these approaches have successfully captured the natural patterns related to urban functions, they have focused only on quantifying single types of urban landscape elements. Multiple sources of urban landscape elements and their interactions have remained undiscussed. As a result, the landscape characteristics that correspond to urban functions are not sufficiently captured.

Based on the previous literature, the necessity to construct effective spatial metrics and functional classification approaches is evident (Jin et al., 2018; Zhang et al., 2019). The limitations and challenges revealed by previous research can be divided into three main aspects. First, since existing spatial metrics can only capture the characteristics of general urban forms or single urban elements, landscape patterns composed of multiple urban elements have not been considered in the classification of urban functions. Thus, new spatial metrics are required for inferring the spatial patterns of these elements. Second, despite the various previously proposed spatial metrics, a sufficient evaluation method is needed to both assess the feasibility of each independent metric and ensure the accuracy of the classified urban functions. Third, integrating quantified landscapes with urban functions and establishing an automatic model for the classification of urban functions with a higher accuracy remain challenging.

Considering these challenges, by measuring urban landscapes using spatial metrics, this paper presents a novel framework for the classification of urban functions. The objectives of this research are to (1) sufficiently characterise urban landscapes by integrating multiple landscape elements using spatial metrics, (2) assess the usability of these spatial metrics in the identification of urban functions, and (3) build an automatic urban function classification model based on the proposed spatial metrics. Based on these objectives, the contributions of this paper are as follows. First, we effectively characterise the landscapes of urban functions. Spatial metrics are designed by measuring multiple landscape elements, including buildings, green land, water bodies, urban facilities, secondary roads and the remaining other areas, and their interactions. Second, the feasibility of these spatial metrics is carefully assessed. We propose a partial correlation analysis to evaluate the sufficiency of each spatial metric in describing urban landscapes. This analysis is followed by an analysis of variance (ANOVA) to ensure the effectiveness of these metrics in indicating specific urban functions. Third, we propose a novel framework to automatically classify urban functions. The assessed spatial metrics are utilised in the classification based on a conditional inference random forest approach. To evaluate the effectiveness of the method, other algorithms, including naïve Bayes, support vector machine (SVM), and random forest, are also considered by comparing their testing accuracies in urban function classification.

The remainder of this paper is organised as follows. Section 2 briefly describes urban landscapes, introduces the experimental data, and discusses the proposed methodology used to classify urban functions. Section 3 demonstrates the performance of our methodology. Section 4 presents a comparison of different algorithms in the classification of urban functions, the improvements in the classification approach by the use of socioeconomic features, and extensive applications of the approach in urban function classification. Section 5 concludes and discusses future directions of our research.

2. Annotation and methodology

2.1. Annotation

To better understand the impact of multiple urban landscape elements on different functions, examples of urban landscape patterns and urban functions are illustrated in Fig. 1. Fig. 1(a), (c) and (e) represent three parcels in high-resolution images, indicating residential, commercial and business, and public and service functions, respectively. The vectorised landscapes of these land parcels are displayed in Fig. 1(b), (d) and (f). Through a visual inspection, the differences between the urban landscape units in the three types of urban functions can be distinguished. The residential functions shown in Fig. 1(a) and (b) display a high density of buildings, a linear distribution of urban facilities among roads, and a lack of water bodies and green land. In particular, urban facilities refer to essential services to the urban environment, such as bus stops, shops and gas stations. Compared to the residential function, the commercial and business functions shown in Fig. 1(c) and (d) exhibit a relatively more even distribution of urban facilities, including buildings and other areas. Fig. 1(e) and (f) reveal that public and service functions are usually occupied by a high percentage of green land or water bodies and far fewer buildings. The above comparisons show the potential for classifying urban functions through the quantification of multiple urban landscape elements.

With respect to measuring urban landscapes, we propose spatial metrics to quantify various urban landscape elements. The measurement is based on the hypothesis shown in Fig. 2. The landscapes and their elements are illustrated in Fig. 2(a). On this basis, the measurements of urban landscapes are illustrated in Fig. 2(b). Each unique urban landscape parcel is divided using main roads. Thus, the elements in the landscapes are related to urban landscape elements, such as buildings, water bodies and green land. Through the utilisation of these urban landscape elements, we aim to capture, evaluate, and distinguish landscape characteristics for use in classifying urban functions.

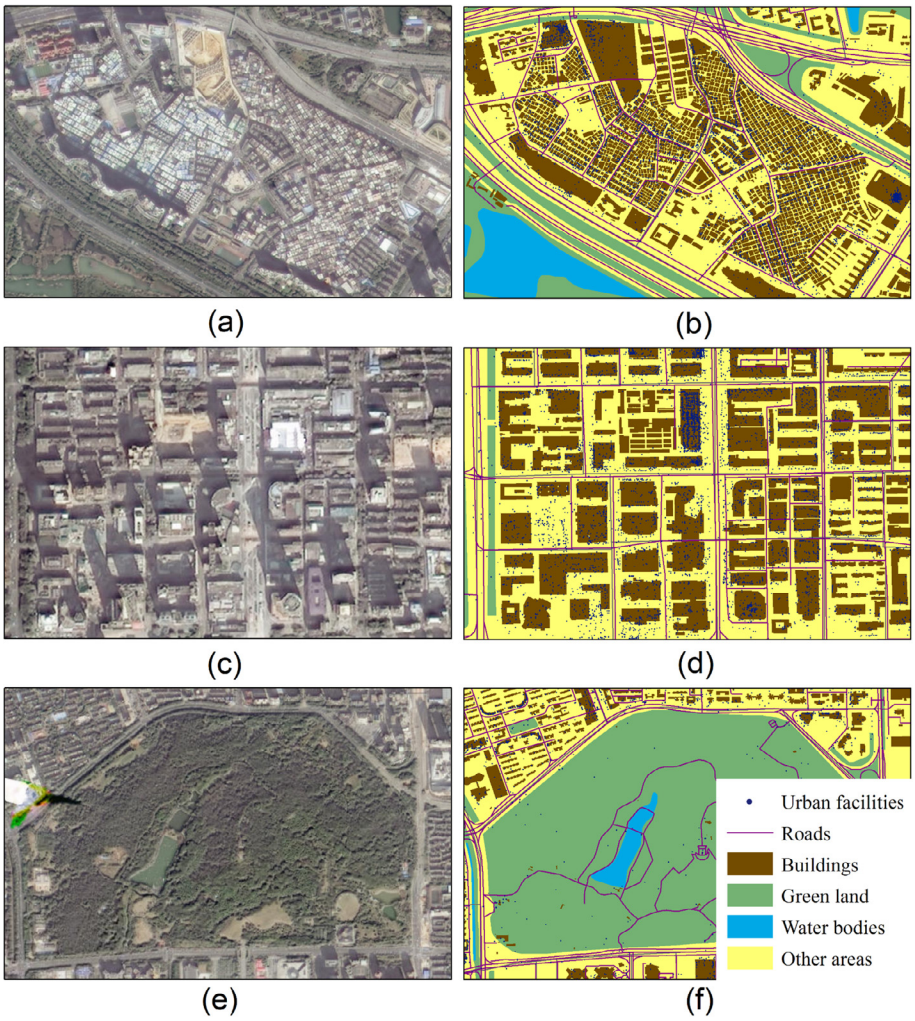


Fig. 1. Different urban landscapes and their corresponding urban functions. (a) High-resolution image of a residential-based urban landscape. (b) Typical urban landscape of residential functions. (c) High-resolution image of a commercial and business-based urban landscape. (d) Typical urban landscape of commercial and business functions. (e) High-resolution image of a public and service-based urban landscape. (f) Typical urban landscape of public and service functions (high-resolution images are provided by TIANDITU (http://www.tianditu.com/)).

2.2. Methodology

The proposed framework is displayed in Fig. 3. The framework consists of three components, including the illustration of urban landscape elements, the methodology used for the classification of urban functions, and the presented results. Through a brief description of the experimental data, the selection of urban landscape elements is proposed in Section 2.2.1. Then, the methodology is proposed

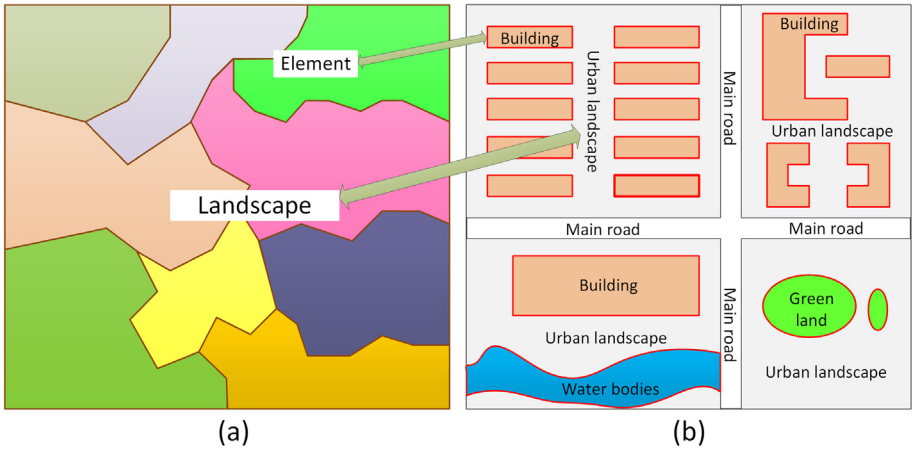


Fig. 2. Illustration of urban landscapes. Elements in landscapes corresponding to urban landscape elements, including buildings, green land, water bodies and other areas. (a) Landscape. (b) Urban landscape.

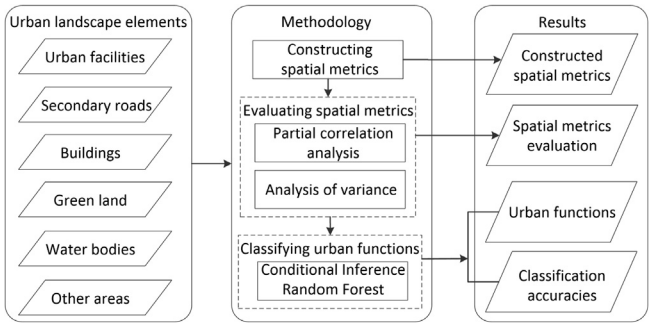


Fig. 3. Overview of the framework.

according to the following three steps: first, spatial metrics are constructed by utilising the selected elements to measure urban landscapes (Section 2.2.2); second, the feasibility of using these spatial metrics in distinguishing different urban functions is evaluated by both a partial correlation analysis and an analysis of variance as described in Section 2.2.3; and third, urban functions are classified by applying the spatial metrics and a conditional inference random forest model (Section 2.2.4). The results are presented in Section 3.

2.2.1. Landscape elements and experimental data

The study area is located in Futian District, Shenzhen, Guangdong Province and occupies a total area of 78.66 km^2. Six types of urban landscape elements, namely, buildings, water bodies, green land, other areas, urban facilities and secondary roads, were chosen to constitute the landscape patterns. In particular, we considered urban landscape parcels land parcels delineated by main roads in which the landscape elements are assembled to represent different landscape environments. Subsequently, the proposed elements were utilised to classify each parcel into one specific urban function, namely, residential, commercial and business, or public and service function.

Fig. 4(a) displays the spatial distribution of the three types of urban landscape elements, including buildings, green land, and water bodies. In particular, the building datasets were obtained from the Urban Planning, Land and Resources Commission of Shenzhen Municipality (Municipality Oceanic Administration of Shenzhen), while the green land and water body polygons were collected via AutoNavi (www.amap.com). The buildings, green land and water bodies occupy areas of 10.60 km^2, 14.52 km^2, and 2.49 km^2, respectively, constituting approximately 13.48%, 18.46% and 3.17% of the study area. The remaining regions are regarded as a single urban landscape element type, namely, other areas.

Fig. 4(b) shows the other two urban landscape elements, namely, urban facilities and secondary roads, and urban landscape parcels. Here, urban facilities are represented by a total of 139,831 points of interest (POIs). Non-urban facility-related POIs (such as the names of administrative and natural areas) are excluded from this study. Secondary roads were extracted from all types of roads as one type of element, whereas main roads were utilised to delineate urban landscape parcels. Both POIs and secondary roads were obtained through AutoNavi.

Based on the above experimental data, detailed descriptions of the urban landscape are further presented in Fig. 5. This figure highlights an urban landscape parcel exhibiting an urban landscape pattern and consisting of basic urban landscape elements. The following section proposes spatial metrics for measuring urban landscapes.

2.2.2. Constructing spatial metrics for the measurement of urban landscapes

In this section, spatial metrics are proposed for the quantification of the characteristics of urban landscapes. With the primary aim of distinguishing the functions in each urban landscape parcel, we construct landscape-level spatial metrics to investigate the spatial distribution

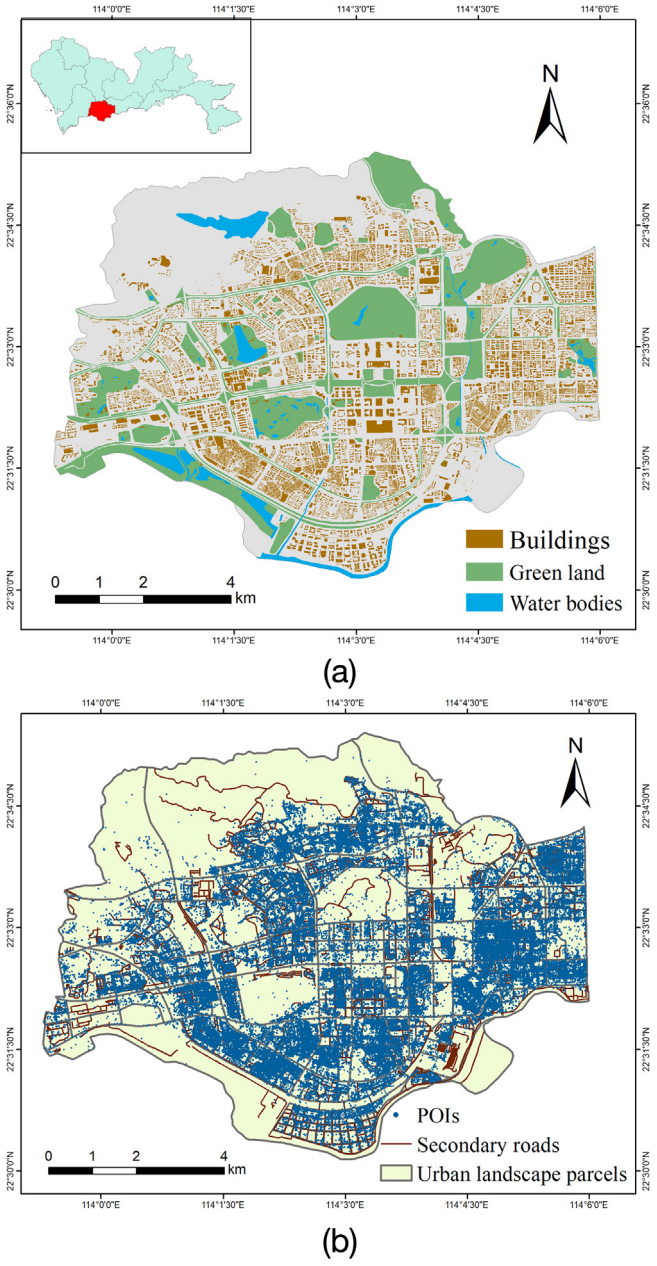


Fig. 4. Study area and experimental data. (a) Spatial distribution of the buildings, green land and water bodies in the study area. (b) POIs and secondary roads in the study area. Urban landscape parcels delineated through main roads.

and patterns among different landscape elements. The construction of each spatial metric is based on a specific purpose with respect to measuring the variation in urban landscape distributions, and thus, these metrics are either modified based on conventional landscape metrics or designed to delineate particular urban landscape patterns.

Eight types of spatial metrics are constructed from different perspectives. The mathematical details of the metrics, including their abbreviations, formulas and statistical ranges, are summarised in Table 1. To describe the spatial patterns of buildings, we use the building density (BD) as a measure of the separated building area, which is defined as the number of buildings per landscape parcel. The BD can be considered an effective indicator for distinguishing public places with few buildings from other places with high building densities. Increased BD values usually indicate well-developed urban areas, such as intensive residential areas. Furthermore, we apply the average building height



Fig. 5. Landscape elements comprising the urban landscapes in each parcel, including POIs (urban facilities), secondary roads, buildings, green land, water bodies and other areas.

Table 1
Mathematical details of the spatial metrics constructed for measuring urban landscapes.

Spatial metric	Abbreviation	Formula	Range
Building density	BD	$BD = \frac{N_{i,b}}{A_i}$	$0 \le BD < 1$
Average building height	ABH	$ABH = \frac{\sum_{j=1}^{N_{i,b}} h_{j,b}}{N_{i,b}}$	$0 \le ABH$
Percentage of non-built-up areas	PNA	$PNA = \frac{\sum_{j=1}^{N_{i,g}} a_{j,g} + \sum_{j=1}^{N_{i,w}} a_{j,w}}{A_i}$	$0 \le PNA \le 1$
Ratio of urban facilities to buildings	RFB	$RFB = \frac{N_{i,p}}{\sum_{j=1}^{N_{i,b}} a_{j,b}}$	$0 \le RFB \le 1$
Ratio of urban facilities to other areas	RFO	$RFO = \frac{N_{i,p}}{\sum_{j=1}^{N_{i,o}} a_{j,o}}$	$0 \le RFO \le 1$
Ratio of roads to buildings	RRB	$RRB = \frac{\sum_{j=1}^{N_{i,r}} l_{j,r}}{\sum_{j=1}^{N_{i,b}} a_{j,b}}$	$0 \le RRB \le 1$
Ratio of roads to other areas	RRO	$RRO = \frac{\sum_{j=1}^{N_{i,r}} l_{j,r}}{\sum_{j=1}^{N_{i,o}} a_{j,o}}$	$0 \le RRO \le 1$
Landscape shape index	LSI	$LSI = \frac{E_i}{2\sqrt{\pi A_i}}$	$LSI \le 1$

Note: $N_{i,b}$, $N_{i,g}$, $N_{i,w}$, $N_{i,p}$, $N_{i,o}$ and $N_{i,r}$ denote the total numbers of buildings, green land, water body patches, POIs, other areas and secondary roads in the ith urban landscape, respectively; A_i denotes the total area (m^2) of the ith urban landscape; $h_{j,b}$ denotes the height (m) of the jth building; $a_{j,g}$, $a_{j,w}$, $a_{j,b}$ and $a_{j,o}$ denote the area (m^2) of the jth green land, water body patch, building and other areas, respectively; $l_{j,r}$ refers to the length (m) of the jth secondary road; E_i represents the total edges (m) of the buildings, green land, water bodies, and other areas in the ith urban landscape.

(ABH) by measuring the mean height of the buildings within the same urban landscape parcel. By utilising this metric, we aim to capture the landscape patterns from a 3-dimensional perspective. In particular, buildings with high values of building height are usually located in high-level residential lands and commercial lands. In contrast, many public and service lands contain buildings of much lower heights.

The percentage of non-built-up areas (PNA) indicates the proportion of the total area of both water bodies and green land relative to the entire area of the corresponding urban landscape parcel. Thus, the PNA is utilised to distinguish the dominant types of built-up from the non-built-up areas. For instance, certain public spaces, such as parks, are generally occupied by mostly green land and water bodies.

Regarding the spatial arrangement of the different urban landscape elements, four spatial metrics are proposed. Using the ratio of urban facilities to buildings (RFB) and the ratio of urban facilities to other areas (RFO), we calculate the number of POIs divided by the total area of buildings or other areas within the same urban landscape parcel. Two other metrics include the ratio of roads to buildings (RRB) and the ratio of roads to other areas (RRO). These metrics are calculated by summing the total length of the secondary roads divided by the total area of buildings or other areas. The RFB and RFO reveal the conditions of urban facilities, while the RRB and RRO provide insight into the arrangement of roads relative to buildings and other public regions. These indicators consider the spatial distribution of buildings, urban facilities, roads and other areas evidence of serving different functions. For instance, residential land tends to contain more buildings and less urban facilities than commercial land. Thus, lower values of RFB are usually obtained in residential areas compared with those in commercial land.

In addition, the landscape shape index (LSI) (Patton, 1975; Rutledge, 2003), which calculates the total lengths of the edges of buildings, water bodies, green lands, and other areas, is adopted. The quantification of the LSI is performed to investigate the aggregation and heterogeneity of these urban landscape elements. Generally, the degree of aggregation varies with different urban functions as follows: a value close to 0 indicates a simpler patch distribution, such as shopping malls, whereas higher values represent much more irregular spatial patterns, including high-density residences and attractions.

2.2.3. Evaluating the spatial metrics for urban function classification

To ensure both sufficiency in depicting urban landscapes and effectiveness in distinguishing urban functions, the constructed spatial metrics are evaluated in the following two steps.

First, a partial correlation analysis is proposed to assess the sufficiency of each spatial metric. In particular, the partial correlations between each two spatial metrics are calculated, while the other metrics are considered control variables. Lower correlations close to 0 indicate that the two spatial metrics are more independent of each other, thus suggesting that both metrics can be utilised to capture sufficient urban landscape information. In contrast, correlations that are close to 1 or −1 suggest that the two metrics are highly correlated. As a result, it is not necessary to utilise both metrics because they may represent similar and redundant urban landscape information.

The second step involves an evaluation of the effectiveness of the classification of different urban functions using an ANOVA. The reference urban functions of each urban landscape parcel are required to evaluate the spatial metrics. In the ANOVA, we focus on the F values with respect to the *significance* of each metric. Generally, spatial metrics with lower F values are not strongly related to urban functions. Consequently, the effectiveness of these metrics in classifying urban functions can be ensured.

2.2.4. Classifying urban functions based on conditional inference random forest

Utilising spatial metrics as variables, we focus on the random forest algorithm for the classification of urban function as this ensemble method is more reliable and robust than using individual classifiers (Dietterich, 2000). During the classification process, quantifying the importance of the variables is vital to building the random forest model because variation in the variables can significantly influence the classification results (Archer and Kimes, 2008). The current approaches involve 2 main types of random forest variable importance measures, including Gini importance and permutation importance (Genuer et al., 2010). In a previous study, a random forest model based on Gini importance for urban classification was proposed; however, the inaccuracy of Gini importance revealed by existing research indicates the

need for alternative measures (Strobl et al., 2007).

To fill this gap, the conditional inference random forest model proposed by Strobl et al. (2008) is adopted to improve the performance of the classification of urban functions. To avoid the overestimation of highly related variables, this approach is proposed based on the conditional permutation importance. As a result, conditional distributions of each type of function can be calculated to aid in urban function classification.

Let $\beta^{(t)}$ be the out-of-bag sample for tree t, where $t \in \{1, 2 \cdots\cdots ntree\}$. *ntree* indicates the number of trees in the forest. Then, the conditional permutation importance $VI^{(t)}(X_j)$ of spatial metric X_j in tree t is defined as follows:

$$VI^{(t)}(X_j) = \frac{\sum_{i \in \beta^{(t)}} I(\gamma_i = \hat{\gamma}_i^{(t)})}{|\beta^{(t)}|} - \frac{\sum_{i \in \beta^{(t)}} I(\gamma_i = \hat{\gamma}_{i,\pi j|Z}^{(t)})}{|\beta^{(t)}|} \quad (1)$$

where $\hat{\gamma}_i^{(t)}$ and $\hat{\gamma}_{i,\pi j|Z}^{(t)}$ indicate the classified urban function of the ith observation before and after permutating the value of the spatial metric X_j, respectively. Z refers to all spatial metrics used for the classification of urban functions. Based on the importance of the variable in each tree, the importance $VI(X_j)$ of the spatial metric X_j to the forest is calculated as follows:

$$VI(X_j) = \frac{\sum_{t=1}^{ntree} VI^{(t)}(X_j)}{ntree} \quad (2)$$

When constructing a model for the classification of urban functions based on the conditional permutation importance, two parameters, including *mtry*, which regulates the number of randomly preselected splitting spatial metrics, and *ntree*, must be determined. We adopt a grid search strategy by calculating the accuracies with $mtry \in [1, 10]$ and $ntree \in [200, 2000]$. The analyses of the computed accuracies based on different values of *mtry* and *ntree* are proposed in Section 3.3.

3. Results

3.1. Descriptive and spatial analysis of spatial metrics

Through the quantification of urban landscape elements, the descriptive statistics of the spatial metrics were calculated and are displayed in Table 2. Among the 411 total urban landscape parcels, all 8 spatial metrics contain minimum values close to 0, whereas the variety of maximum values indicates different ranges of the values in these metrics. In addition, except for ABH and LSI, 6 spatial metrics contain similar standard deviation and variance values. The various ABH values indicate that the average building heights significantly change among different urban landscape parcels. These statistics suggest their potential use in the classification of urban function.

From the spatial perspective, the distribution of each quantified spatial metric is shown in Fig. 6. Specifically, according to the metrics proposed in Table 1, eight spatial metrics are calculated based on landscape elements within each urban landscape parcel. Though a visual inspection, one can distinguish the differences in the distributions of the 8 spatial metrics. For instance, compared with the other parcels, the large parcel located in the north of the study area contains much lower values of BD, ABH, RBF, ROF and RRO (see Fig. 6(a), (b), (d), (e) and (g)). Additionally, this parcel contains a medium level of PNA and LSI as presented in Fig. 6(c) and (h), whereas it displays a much higher value of RRB (Fig. 6(f)).

3.2. Spatial metrics evaluation

In this section, the constructed spatial metrics are evaluated in terms of their sufficiency and effectiveness through a partial correlation analysis and ANOVA, respectively. Regarding the efficiency of the spatial metrics, the correlation between each pair of spatial metrics was calculated with the other metrics as variables. Table 3 shows the results of the partial correlations between each of the 8 spatial metrics. Most calculated partial correlations are close to 0, indicating that they are independent and represent different characteristics of urban landscapes.

Notably, the correlation between BD and LSI, namely, 0.715, shows a much higher value of partial correlation. This finding reveals that BD and LSI may represent similar urban landscape features. However, the spatial metrics BD and LSI represent completely different concepts; while BD focuses on delineating the spatial distribution of buildings, LSI indicates the heterogeneity of all types of landscape elements. The similar landscape features indicated may be because both the BD and LSI are highly related to the existing numbers of landscape elements, especially buildings. Since buildings are the dominant element in most landscape parcels (occupying larger areas than the other elements), a high partial correlation value between BD and LSI is acceptable. This issue could be further discussed in future studies examining different urban landscapes.

To further evaluate the effectiveness of the spatial metrics in classifying urban functions, an ANOVA was performed. The normality of all 8 spatial metrics was first examined, and the results displayed normal distributions, which are suitable for an ANOVA. As presented in Table 4, the spatial metrics were calculated with significance equalling 0.000 and F values ranging from 10.230 to 44.775. The results indicate that the values of the spatial metrics vary with respect to the different urban functions and that there is further potential in utilising these metrics for the classification of urban functions.

3.3. Urban function classification

Based on the conditional inference random forest model, the urban functions in the urban landscape parcels were classified using the constructed spatial metrics. We utilised the *cforest* function in R to achieve the modelling. Of a total of 411 urban landscape parcels, 300 parcels were randomly selected as the training data, and the remaining 111 parcels were regarded as testing data. The reference urban functions for both model training and testing were obtained from the Urban Planning, Land & Resources Commission of Shenzhen Municipality (Municipality Oceanic Administration of Shenzhen) (http://www.szgeoinfo.com:5001/msmap/flex/landuse.html) and further

Table 2
Descriptive statistics of the spatial metrics.

Spatial metrics	N	Minimum	Maximum	Mean	Std. Deviation	Variance
BD	411	0.00000	0.00273	0.00030	0.00038553	0.000
ABH	411	0.00000	195.000	29.5534	24.84104683	617.078
PNA	411	0.00000	0.89766	0.07590	0.15183816	0.023
RFB	411	0.00000	0.13029	0.00835	0.01151707	0.000
RFO	411	0.00003	0.02648	0.00233	0.00266159	0.000
RRB	411	0.00000	0.75472	0.03584	0.06769773	0.005
RRO	411	0.00000	0.05206	0.00958	0.00901437	0.000
LSI	411	1.12911	23.2039	4.69482	2.91737108	8.511
Valid N (listwise)	411					

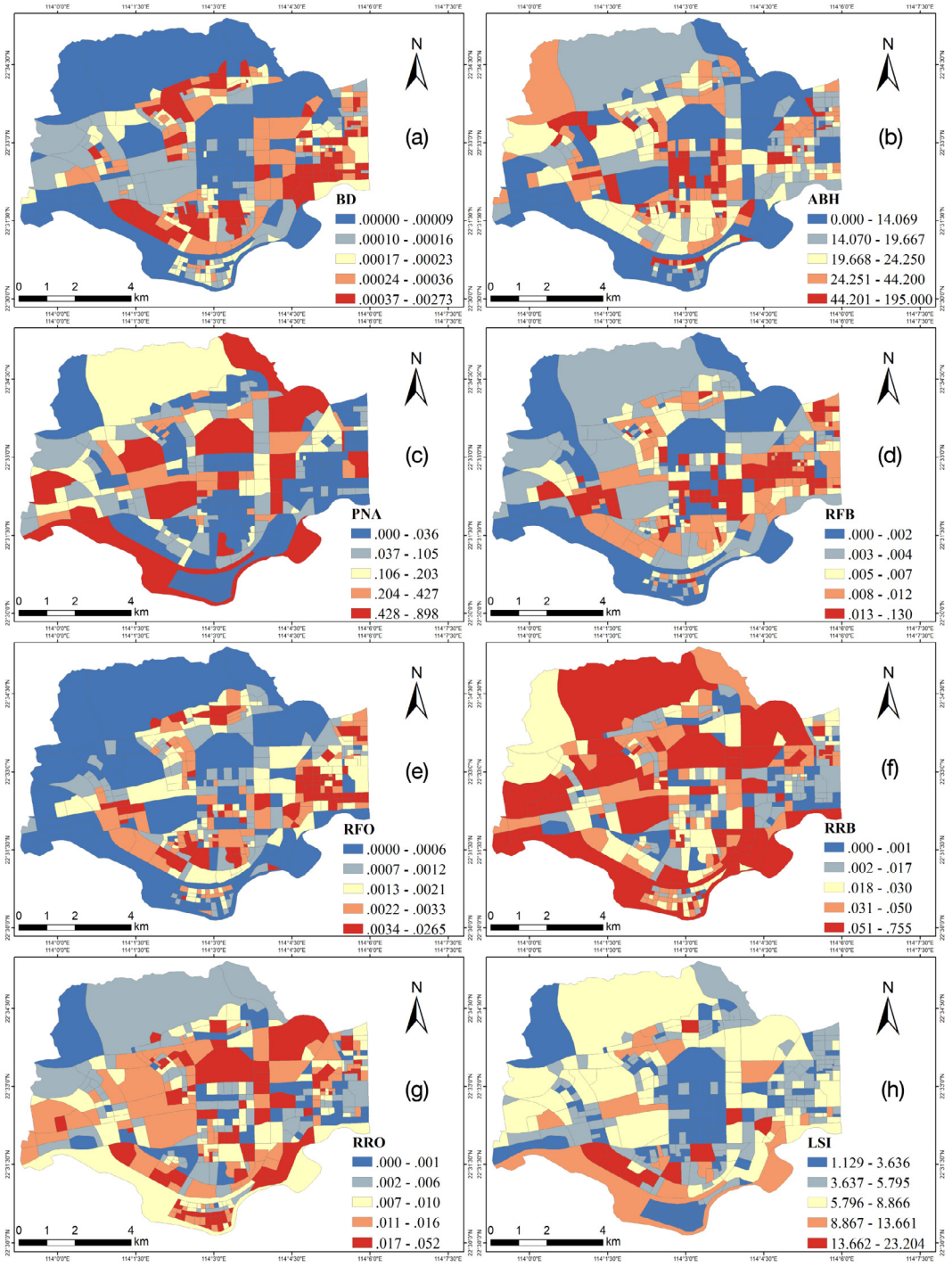


Fig. 6. Spatial distributions of the calculated spatial metrics. (a) Spatial distribution of BD; (b) Spatial distribution of ABH; (c) Spatial distribution of PNA; (d) Spatial distribution of RBF; (e) Spatial distribution of RFO; (f) Spatial distribution of RRB; (g) Spatial distribution of RRO; (h) Spatial distribution of LSI.

reclassified to match the proposed nomenclature. In particular, the importance score of each spatial metric was calculated based on the conditional permutation importance and is shown in Table 5.

Utilising the importance scores as measurements, the changes in testing accuracies were investigated. In particular, a grid search strategy was proposed to choose parameters *mtry* and *ntree*. Values with *mtry* ∈ [1, 10] and *ntree* ∈ [200, 2000] were combined to measure the differences in classification accuracies. A confusion matrix was first proposed (shown in Fig. 7) to determine the ranges of *mtry*. On this basis, each value of *mtry* was combined with *ntree* ∈ [200, 2000] to calculate the specific classification accuracies (shown in Fig. 8).

As shown in Fig. 7, with the parameters *mtry* and *ntree* changing within [1, 10] and [200, 2000], respectively, the testing accuracies vary from approximately 0.71 to 0.77 with darker blue colours representing higher values. Since urban functions reflect both physical patterns and socioeconomic features (Liu et al., 2015), the relatively low testing accuracies based solely on landscape features are acceptable to some extent. This issue is further discussed in Section 4.2. In particular, the accuracies based on the parameter *mtry* ∈ [1, 4] are much lower, with no significant changes, whereas using *mtry* ∈ [5, 10], the accuracies vary with the different values of the parameter *ntree*. Accordingly, to effectively determine the values of the parameters, we specifically

Table 3
Partial correlations between the spatial metrics.

	BD	ABH	PNA	RFB	RFO	RRB	RRO	LSI
BD	1	−0.175	−0.295	−0.039	0.259	−0.048	−0.013	0.715
Significance	–	0.000	0.000	0.437	0.000	0.331	0.794	0.000
ABH	−0.175	1	−0.1	0.166	0.13	−0.074	−0.004	−0.009
Significance	0.000	–	0.045	0.001	0.009	0.138	0.933	0.849
PNA	−0.295	−0.1	1	0.023	−0.078	0.228	−0.15	0.28
Significance	0.000	0.045	–	0.645	0.116	0.000	0.002	0.000
RFB	−0.039	0.166	0.023	1	0.475	−0.029	−0.05	0.005
Significance	0.437	0.001	0.645	–	0.000	0.562	0.318	0.918
RFO	0.259	0.13	−0.078	0.475	1	−0.053	0.005	−0.102
Significance	0.000	0.009	0.116	0.000	–	0.287	0.920	0.040
RRB	−0.048	−0.074	0.228	−0.029	−0.053	1	0.354	0.003
Significance	0.331	0.138	0.000	0.562	0.287	–	0.000	0.952
RRO	−0.013	−0.004	−0.15	−0.05	0.005	0.354	1	0.027
Significance	0.794	0.933	0.002	0.318	0.920	0.000	–	0.583
LSI	0.715	−0.009	0.28	0.005	−0.102	0.003	0.027	1
Significance	0.000	0.849	0.000	0.918	0.040	0.952	0.583	–

Table 4
ANOVA results of the spatial metrics.

Spatial metric	F	Sig.
BD	33.693	0.000
ABH	12.636	0.000
PNA	44.775	0.000
RFB	35.160	0.000
RFO	22.396	0.000
RRB	12.441	0.000
RRO	10.230	0.000
LSI	28.586	0.000

Table 5
Conditional permutation importance of the spatial metrics.

Spatial metric	Importance score
BD	0.01762
ABH	0.03460
PNA	0.01013
RFB	0.04693
RFO	0.00119
RRB	0.00042
RRO	0.00027
LSI	0.01013

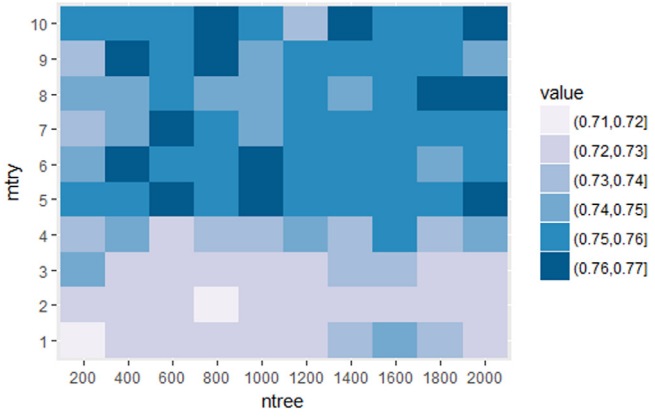


Fig. 7. Testing accuracies of a grid search using the parameters *mtry* ∈ [1, 10] and *ntree* ∈ (200, 400, 600……2000).

focused on the patterns of the testing accuracy with *mtry* ∈ [5, 10] and *ntree* ∈ [200, 2000].

Fig. 8 presents the testing accuracy patterns using *mtry* ∈ [5, 10] and *ntree* ∈ [200, 2000]. Regarding each selected value of *mtry*, the accuracies change irregularly as *ntree* increases. Generally, higher values of *ntree* may lead to model overfitting and further result in biases in the classification of urban function. Thus, the parameters *mtry* and *ntree* should be selected with the corresponding accuracies stabilised and not significantly changed. In particular, as displayed in Fig. 8(a), (b), (d) and (f), when *mtry* ∈ (5, 6, 8, 10), the testing accuracies increase greatly as *ntree* reaches 2000. Meanwhile, Fig. 8(e) displays a dramatic decrease in the testing accuracy with *mtry* = 9 and *tree* increasing to 2000. A more stabilised tendency of accuracy is observed in Fig. 8(c) with *mtry* = 7 and *ntree* increasing from 1200 to 2000. This finding indicates that *mtry* = 7 is appropriate for reducing model overfitting regardless of the increasing value of *ntree*. As a result, the urban functions were classified using the conditional inference random forest model with *mtry* = 7 and *ntree* = 1200.

Utilising the conditional inference random forest model with the selected parameters, the conditional distributions of the three types of urban functions are shown in Fig. 9. By calculating the classification probabilities, one can distinguish among the urban functions with various levels of confidence to ensure better classification accuracy. Here, the urban function in each parcel is determined according to the highest value of the probabilities. Based on this calculation, the classification results are displayed in Fig. 10(a). For a visual inspection, the corresponding reference urban functions are shown in Fig. 10(b).

According to the validation using reference data, the classification accuracy of the total 411 urban landscape parcels is displayed in Table 6. An overall accuracy of 0.813 is obtained, which is close to the testing accuracy of 0.757 (*mtry* = 7 and *ntree* = 1200). Regarding the producer's accuracies, the highest value of 0.92 is found for the residential function, whereas the commercial and business function shows an accuracy of 0.831. In addition, the classification of the public and service function has the lowest accuracy at only 0.453. Regarding the user accuracy, the residential, commercial and business and public and service functions achieve similar accuracy values of 0.819, 0.801 and 0.813, respectively.

4. Discussion

4.1. Comparing different algorithms for urban function classification

To evaluate the accuracy of urban classifications using the conditional inference random forest model, three other types of algorithms, namely, naïve Bayes (Maron, 1961), SVM (Cortes and Vapnik, 1995) and random forest (Ho, 1995), were adopted to classify the urban landscape parcels in the study area. These algorithms are known as popular machine learning methods that show advanced classification effectiveness when analysing large datasets (Witten et al., 2016). In

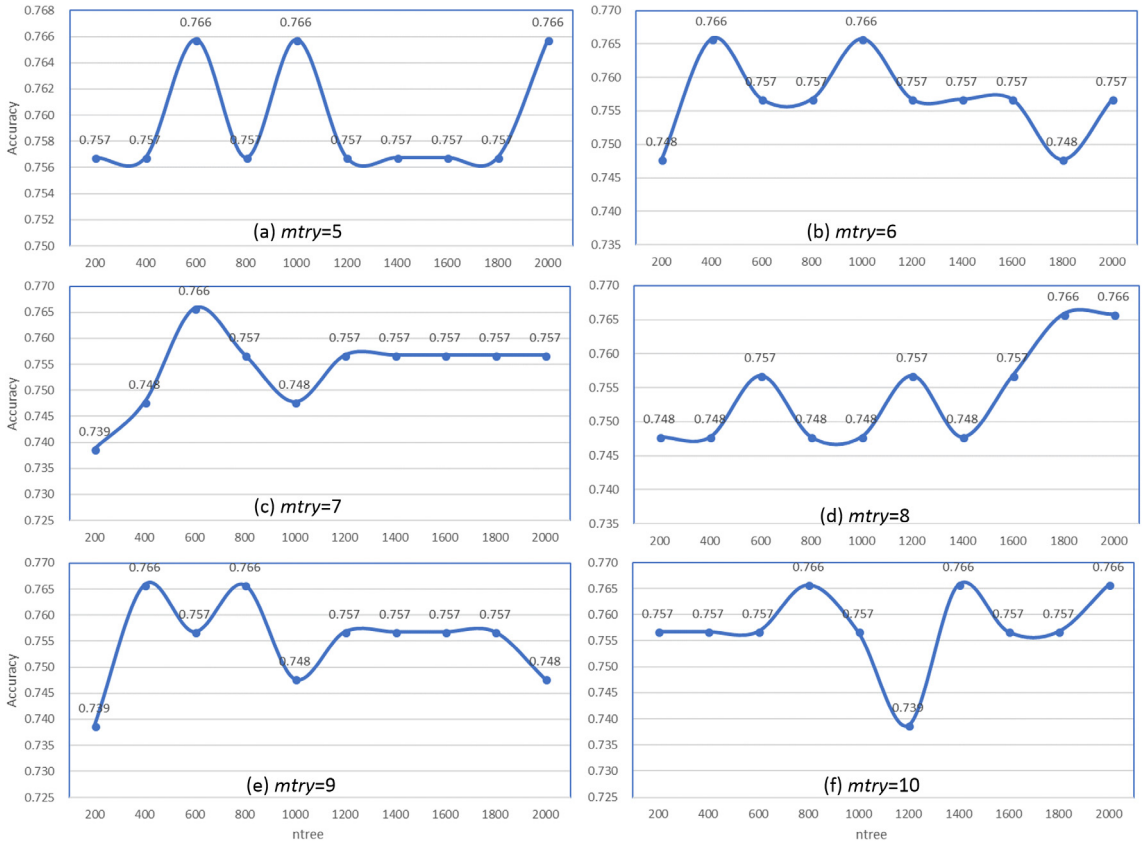


Fig. 8. Testing accuracy patterns using *mtry* ∈ [5, 10] and *ntree* ∈ [200, 2000]. (a) *mtry* = 5; (b) *mtry* = 6; (c) *mtry* = 7; (d) *mtry* = 8; (e) *mtry* = 9; (f) *mtry* = 10.

particular, the parameters *Laplace* in naïve Bayes, *gamma* and *cost* in SVM, and *mtry* and *ntree* in random forest were estimated and further determined using a grid search method. In addition, for the SVM algorithm, a sigmoid kernel was chosen to build the urban function classifier. For the random forest algorithm, the variables were measured based on the Gini importance, which can be compared with the conditional permutation importance used in this study.

Table 7 displays the testing accuracies of the classification of urban functions using the four different algorithms. The conditional inference random forest algorithm achieves the highest overall accuracy of 0.757, whereas the testing accuracies of naïve Bayes, SVM and random forest are 0.667, 0.595 and 0.712, respectively. The overall accuracies of the random forest and the conditional inference random forest (0.712 and 0.757, respectively) also show an advantage of conditional permutation importance over Gini importance. For the producer's and user's accuracies of the residential functions, the respective values of 0.652 and 0.664 are obtained in naïve Bayes, 0.739 and 0.618 in SVM, and 0.826 and 0.691 in random forest. The conditional inference random forest approach still achieves the highest accuracy with values of 0.913 and 0.737. Regarding the accuracies of the commercial and business function, SVM and random forest achieve accuracies of approximately 0.5 and 0.7, respectively. Naïve Bayes achieves accuracies of 0.63 and 0.75, while the conditional inference random forest yields much better results of 0.804 and 0.755. Regarding the public and service function, significant divergence is observed between the producer's and user's accuracies. While the user's accuracies in SVM, random forest and conditional inference random forest are 0.556, 0.857 and 1, respectively, the producer's accuracies are much lower at only 0.263, 0.316 and 0.263, respectively. Despite the variation in the testing accuracies, the proposed comparative experiment confirms the advantages of using the conditional inference random forest for classification.

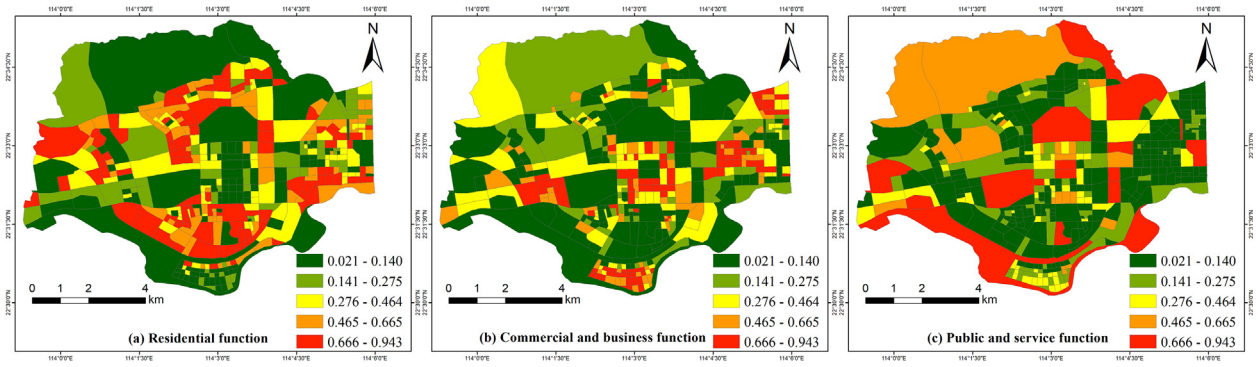


Fig. 9. Conditional distributions of the urban functions. (a) Residential function; (b) Commercial and business function; (c) Public and service function.

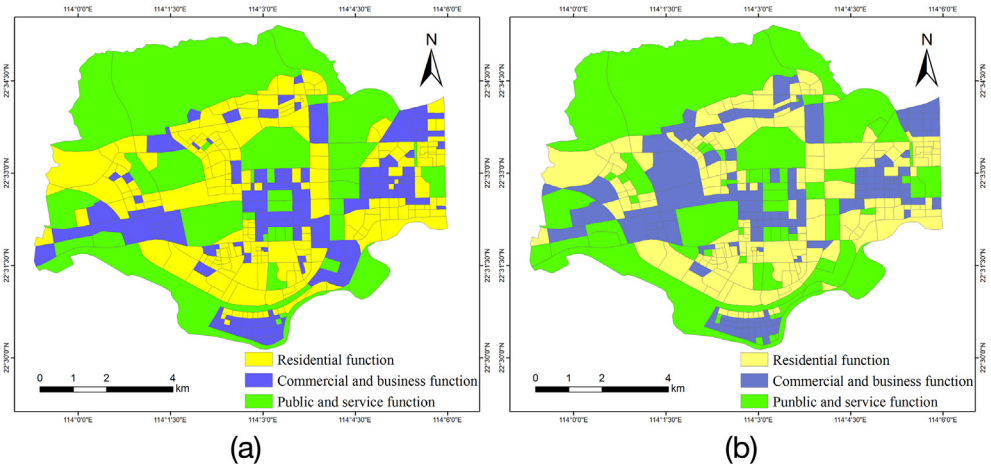


Fig. 10. Classification of the urban functions using spatial metrics and a comparison with reference data. (a) Urban function classification; (b) Reference data (derived and reclassified from land use data provided by the Urban Planning, Land Resources Commission of Shenzhen Municipality (Municipality Oceanic Administration of Shenzhen)).

4.2. Improving urban function classification accuracy with socioeconomic features

In general, urban functions reflect both physical characteristics and socioeconomic features (Liu et al., 2015; Xing and Meng, 2018). As this research mainly focuses on capturing landscape information in cities, social patterns, which are usually related to human activities, are ignored. The impact of lacking socioeconomic features has been revealed in this study in the following aspects. First, as shown in Tables 6 and 7, the accuracy of the classification of the urban functions and the testing accuracy are 0.813 and 0.757, respectively. These relatively low accuracies indicate that incorporating socioeconomic features is essential for improving the classification accuracy. Second, only three types of urban functions, namely, residential, commercial and business, and public and service functions, are included in this study. Through the utilisation of social information, more urban functions can be classified. For instance, one could delineate the commercial and business functions independently or even detect shopping malls within commercial areas.

To effectively capture socioeconomic features, the potential use of crowdsourced data can be considered to assist in the classification of urban functions. Generally, various types of crowdsourced data are available to represent human activity patterns and, thus, provide an effective approach to investigate urban function-related social information. Much research has been conducted to obtain urban function characteristics using crowdsourced data. Frias-Martinez and Frias-Martinez (2014) analysed the temporal patterns of Twitter check-in data and utilised these regularities for the detection of residential, business and industrial functions. In addition, the potential use of other data sources has been employed, including Tencent user density (Chen et al., 2017), mobile phone data (Pei et al., 2014), metro smart card records (Gong et al., 2017), and trajectory data (Liu et al., 2012).

Regarding the exploration of crowdsourced data, the approach proposed in this study can be improved by integrating both physical and social information. Existing studies have investigated the combination of these two perspectives by extracting physical characteristics from remotely sensed images and socioeconomic features from human activities (Hu et al., 2016; Liu et al., 2017b; Tu et al., 2018; Zhang et al., 2017). Despite these attempts, two important issues remain. First, the generated physical characteristics reveal the whole picture or coarse landscapes of only the urban scene and, thus, fail to capture the detailed landscape patterns presented in this study. Second, the integration of detailed landscape patterns and socioeconomic features is challenging. These two challenges will be investigated in a future study.

Table 7
Testing the accuracies of urban function classification based on different algorithms (PA refers to producer's accuracy, UA refers to user's accuracy, and OA refers to overall accuracy).

		Residential function	Commercial and business function	Public and service function	OA
Naïve Bayes	PA	0.652	0.63	0.789	0.667
	UA	0.644	0.75	0.577	
SVM	PA	0.739	0.587	0.263	0.595
	UA	0.618	0.574	0.556	
Random forest	PA	0.826	0.761	0.316	0.712
	UA	0.691	0.714	0.857	
Conditional inference random forest	PA	0.913	0.804	0.263	0.757
	UA	0.737	0.755	1	

4.3. Extending applications in urban function classification by measuring urban landscapes

Regarding the effectiveness of depicting landscape patterns for urban function classification, the proposed approaches should be considered for improvement for use in extensive applications. The first issue is modifying the model to fit more types of urban functions. The current approach focuses on only a few functions, including residential, commercial and business, and public and services. To some extent, this approach limits city planners and decision-makers in making full use of this framework in applications. In fact, the status of urban functions is more complex and far beyond conceptual (Wu et al., 2018). According to the existing governmental land use data in China, more specific

Table 6
Urban function classification accuracies using the parameters *mtry* = 7 and *ntree* = 1200 (PA refers to producer's accuracy, UA refers to user's accuracy).

	Residential function	Commercial and business function	Public and service function	Total	UA
Residential function	172	23	15	210	**0.819**
Commercial and business function	13	133	20	166	**0.801**
Public and service function	2	4	29	35	**0.829**
Total	187	160	64	411	
PA	**0.920**	**0.831**	**0.453**		**0.813**

functions should be included in the nomenclature. For instance, public and services can be divided into public facilities, green space and squares, administration and public services (Yao et al., 2017). In addition, many multifunctional parcels exist in urban areas, such as in the integration of residential and commercial functions (Yue et al., 2017). Faced with these problems, it is necessary to improve the ability to depict urban landscapes for extensive urban function classification.

In addition to increasing the urban function types, choosing alternative data sources for function classification is another issue remaining to be solved. Although the vectorised data sources used in this study including buildings, water and green land have been proven to be effective for urban function classification, they are quite difficult to obtain in some cities (Xing and Meng, 2018). Due to the unavailability of data sources, it is necessary to find data sources that are easy to obtain. On the one hand, the freely available data from OpenStreetMap can be utilised to delineate urban landscapes, such as buildings and natural polygons (Haklay and Weber, 2008). On the other hand, high-resolution images have the potential to be processed to obtain the necessary vectorised urban data. These images are available on many platforms, such as Google Earth and MAPWORLD, making it possible to utilise these images for extensive applications of the proposed approach.

5. Conclusion

With the increasing processes of urban forms, the corresponding landscapes hold great potential for use in understanding the classification of the various distributions of urban functions. This study proposed a novel framework for urban function classification by measuring fine-scale urban landscapes using spatial metrics. First, using 6 urban landscape elements (urban facilities, secondary roads, buildings, green land, water bodies and other areas), 8 types of spatial metrics were constructed to capture urban landscape patterns from various perspectives. Then, the constructed spatial metrics were evaluated in terms of both their sufficiency in delineating landscape information and their effectiveness in classifying urban functions using a partial correlation analysis and an ANOVA, respectively. Finally, for the classification of urban functions, a conditional inference random forest method was adopted using the spatial metrics. The results based on the experimental data reveal testing and overall classification accuracies of 0.757 and 0.813, respectively. In addition, testing accuracies of 0.667, 0.595 and 0.712 were obtained in comparative experiments using naïve Bayes, SVM and random forest, respectively. These findings indicate the effectiveness of the conditional inference random forest and further ensure the reliability of the proposed framework for urban function classification.

Based on the proposed approach, future studies investigating the classification of urban functions could focus on the following two aspects. On the one hand, regarding the evaluation of the spatial metrics discussed in Section 2.2.3, better approaches to selecting metrics are required for filtering ineffective metrics and better capturing urban landscape patterns. On the other hand, to improve the accuracy of the classification of urban functions, the integration of socioeconomic features with landscape patterns could be considered.

Declaration of Competing Interest

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

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant Nos. 41971406; 41501420), Shandong Science and Technology Research Project (Grant No. 2017GSF22102), Natural Science Foundation of Shandong Province of China (Grant No. ZR2019MD043), National Keypoint Research and Invention Program of Shandong Province (Grant No. 2017CXGC0304-X).

References

Abrantes, P., Rocha, J., Marques da Costa, E., Gomes, E., Morgado, P., Costa, N., 2017. Modelling urban form: A multidimensional typology of urban occupation for spatial analysis. Environment and Planning B: Urban Analytics and City Science, 2399808317700140.
Alberti, M., Marzluff, J.M., 2004. Ecological resilience in urban ecosystems: linking urban patterns to human and ecological functions. Urban Ecosyst. 7, 241–265.
Anderson, J.R., 1976. A land use and land cover classification system for use with remote sensor data. US Government Printing Office.
Archer, K.J., Kimes, R.V., 2008. Empirical characterization of random forest variable importance measures. Comput. Stat. Data Anal. 52, 2249–2260.
Barnsley, M.J., Barr, S.L., 1997. Distinguishing urban land-use categories in fine spatial resolution land-cover data using a graph-based, structural pattern recognition system. Comput. Environ. Urban Syst. 21, 209–225.
Barr, S., Barnsley, M., 1997. A region-based, graph-theoretic data model for the inference of second-order thematic information from remotely-sensed images. Int. J. Geograph. Informat. Sci. 11, 555–576.
Baus, P., Kováč, U., Pauditšová, E., Kohutková, I., Komorník, J., 2014. Identification of interconnections between landscape pattern and urban dynamics—Case study Bratislava, Slovakia. Ecol. Ind. 42, 104–111.
Bertaud, A., Renaud, B., 1997. Socialist cities without land markets. J. Urban Econom. 41, 137–151.
Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., Xu, X., Pei, F., 2017. Delineating urban functional areas with building-level social media data: a dynamic time warping (DTW) distance based k-medoids method. Landscape Urban Plann. 160, 48–60.
Chen, Z., Xu, B., Devereux, B., 2014. Urban landscape pattern analysis based on 3D landscape models. Appl. Geogr. 55, 82–91.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learn. 20, 273–297.
Dietterich, T.G., 2000. Ensemble Methods in Machine Learning, International Workshop on Multiple Classifier Systems. Springer, pp. 1–15.
Frias-Martinez, V., Frias-Martinez, E., 2014. Spectral clustering for sensing urban land use using Twitter activity. Eng. Appl. Artif. Intell. 35, 237–245.
Genuer, R., Poggi, J.-M., Tuleau-Malot, C., 2010. Variable selection using random forests. Pattern Recogn. Lett. 31, 2225–2236.
Gong, Y., Lin, Y., Duan, Z., 2017. Exploring the spatiotemporal structure of dynamic urban space using metro smart card records. Comput. Environ. Urban Syst. 64, 169–183.
Haklay, M., Weber, P., 2008. Openstreetmap: user-generated street maps. Ieee Pervas Comput 7, 12–18.
Hermosilla, T., Palomar-Vázquez, J., Balaguer-Beser, Á., Balsa-Barreiro, J., Ruiz, L.A., 2014. Using street based metrics to characterize urban typologies. Comput. Environ. Urban Syst. 44, 68–79.
Ho, T.K., 1995. Random decision forests, Proceedings of 3rd international conference on document analysis and recognition. IEEE 278–282.
Hu, T., Yang, J., Li, X., Gong, P., 2016. Mapping urban land use by using landsat images and open social data. Remote Sensing 8, 151.
Jin, R., Gong, J., Deng, M., Wan, Y., Yang, W., 2018. A spatial-adjacency-based approach for analyzing urban landscape structure. Trans. GIS 22, 1649–1672.
Kane, K., Connors, J.P., Galletti, C.S., 2014. Beyond fragmentation at the fringe: a path-dependent, high-resolution analysis of urban land cover in Phoenix, Arizona. Appl. Geogr. 52, 123–134.
Li, H., Peng, J., Yanxu, L., Yi'na, H, 2017. Urbanization impact on landscape patterns in Beijing City, China: a spatial heterogeneity perspective. Ecol. Ind. 82, 50–60.
Liu, M., Hu, Y.-M., Li, C.-L., 2017a. Landscape metrics for three-dimensional urban building pattern recognition. Appl. Geogr. 87, 66–72.
Liu, T., Yang, X., 2015. Monitoring land changes in an urban area using satellite imagery, GIS and landscape metrics. Appl. Geogr. 56, 42–54.
Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., Hong, Y., 2017b. Classifying urban land use by integrating remote sensing and social media data. Int. J. Geograph. Informat. Sci. 1–22.
Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., Shi, L., 2015. Social sensing: a new approach to understanding our socioeconomic environments. Ann. Assoc. Am. Geogr. 105, 512–530.
Liu, Y., Wang, F., Xiao, Y., Gao, S., 2012. Urban land uses and traffic 'source-sink areas': evidence from GPS-enabled taxi data in Shanghai. Landscape Urban Plann. 106, 73–87.
Liu, Y., Zhang, X., Kong, X., Wang, R., Chen, L., 2018. Identifying the relationship between urban land expansion and human activities in the Yangtze River Economic Belt, China. Appl. Geogr. 94, 163–177.
Manonmani, R., Suganya, G.M.D., 2010. Remote sensing and GIS application in change detection study in urban zone using multi temporal satellite. Int. J Geomat. Geosci. 1, 60.
Maron, M.E., 1961. Automatic indexing: an experimental inquiry. J. ACM (JACM) 8, 404–417.
Matsuoka, R.H., Kaplan, R., 2008. People needs in the urban landscape: analysis of landscape and urban planning contributions. Landsc. Urban Plann. 84, 7–19.
McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002. FRAGSTATS: spatial pattern analysis program for categorical maps.
Montanges, A.P., Moser, G., Taubenböck, H., Wurm, M., Tuia, D., 2015. Classification of urban structural types with multisource data and structured models, Urban Remote

Sensing Event (JURSE), 2015 Joint. IEEE 1–4.
Patton, D.R., 1975. A diversity index for quantifying habitat“ edge”. Wildl. Soc. Bull. 1973–2006 (3), 171–173.
Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T., Zhou, C., 2014. A new insight into land use classification based on aggregated mobile phone data. Int. J. Geograph. Inform. Sci. 28, 1988–2007.
Rutledge, D.T., 2003. Landscape indices as measures of the effects of fragmentation: can pattern reflect process?.
Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinform. 9, 307.
Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinform. 8, 25.
Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., Dech, S., 2014. New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. Appl. Geogr. 47, 137–153.
Tsai, Y.-H., 2005. Quantifying urban form: compactness versus' sprawl'. Urban Studies 42, 141–161.
Tu, W., Hu, Z., Li, L., Cao, J., Jiang, J., Li, Q., Li, Q., 2018. Portraying urban functional zones by coupling remote sensing imagery and human sensing data. Remote Sens. 10, 141.
Van de Voorde, T., Jacquet, W., Canters, F., 2011. Mapping form and function in urban areas: an approach based on urban metrics and continuous impervious surface data. Landscape Urban Plann. 102, 143–155.
Vanderhaegen, S., Canters, F., 2017. Mapping urban form and function at city block level using spatial metrics. Landscape Urban Plann. 167, 399–409.
Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. Remote Sens. Environ. 117, 34–49.
Witten, I.H., Frank, E., Hall, M.A., Pal, C.J., 2016. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann.
Wu, L., Cheng, X., Kang, C., Zhu, D., Huang, Z., Liu, Y., 2018. A framework for mixed-use decomposition based on temporal activity signatures extracted from big geo-data. Int. J. Digital Earth 1–19.
Xing, H., Meng, Y., 2018. Integrating landscape metrics and socioeconomic features for urban functional region classification. Comput. Environ. Urban Syst. 72, 134–145.
Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., Mai, K., 2017. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. Int. J. Geograph. Informat. Sci. 31, 825–848.
Yue, Y., Zhuang, Y., Yeh, A.G., Xie, J.-Y., Ma, C.-L., Li, Q.-Q., 2017. Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. Int. J. Geograph. Inform. Sci. 31, 658–675.
Zhang, A., Xia, C., Chu, J., Lin, J., Li, W., Wu, J., 2019. Portraying urban landscape: a quantitative analysis system applied in fifteen metropolises in China. Sustain. Cit. Soc. 46 101396.
Zhang, X., Du, S., Wang, Q., 2017. Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data. ISPRS J. Photogramm. Remote Sens. 132, 170–184.
Zheng, Y., Capra, L., Wolfson, O., Yang, H., 2014. Urban computing: concepts, methodologies, and applications. ACM Trans. Intellig. Syst. Technol. (TIST) 5, 38.