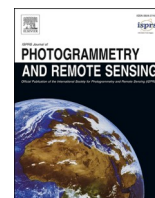


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

ISPRS Journal of Photogrammetry and Remote Sensing

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

Urban functional zone mapping by integrating high spatial resolution nighttime light and daytime multi-view imagery

Xin Huang^{a,b}, Junjing Yang^a, Jiayi Li^{a,*}, Dawei Wen^c

^a School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, PR China

^b State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, PR China

^c College of Public Administration, Huazhong Agricultural University, Wuhan 430070, PR China

ARTICLE INFO

Keywords:

Multi-view
High-resolution
Urban
Nighttime light
Urban functional zones
Multi-source
Data fusion

ABSTRACT

Functional zones are the basic units of cities and the mapping work is fundamental to urban management, investigation, and research. The existing urban functional zone (UFZ) mapping methods usually utilize visual features from high spatial resolution optical images, and focus on two-dimensional image features, such as texture and landscape. However, UFZ is a comprehensive concept including geographical, social, and economic aspects. Therefore, it is reasonable to simultaneously take into account the characteristics of both human activities and image visual features for its accurate interpretation. Multi-view optical satellite images can delineate the physical characteristics of a city, i.e., 3D structures; on the other hand, high spatial resolution nighttime light images are important signals of human activities. These two data sources can be complementary in representing urban landscapes. However, to our knowledge, in the current literature, neither multi-view images nor high spatial resolution nighttime light images have been used for UFZ mapping, and it is poorly understood whether their individual or combined use can achieve satisfactory results. Therefore, in this study, a daytime and nighttime data fusion method, that is, the fusion of daytime multi-view optical images (Ziyuan3-01, 2.1 m) and high spatial resolution nighttime light images (Jilin1-07, 0.92 m), was proposed for UFZ mapping. In particular, a building enhanced nighttime light index (BENI) was proposed to improve the ability of the nighttime light images in discriminating between different functional zones. To verify the effectiveness of the proposed method, experiments were conducted in two megacities of China, i.e., Beijing and Wuhan. Our results indicated that: 1) an OA of ~80% was obtained by the spectral based method; 2) the addition of multi-view features led to ~84% OA, an increment of 4% compared with the spectral features; and 3) the inclusion of nighttime features achieved 85–90% OA, which further improved the OA of daytime features by 1–6%. It was also shown that the proposed BENI was superior to the original nighttime light brightness in identifying functional zones. In general, this study verified the effectiveness and complementarity of daytime (including multispectral and multi-view images) and nighttime images in UFZ mapping, and provided new thoughts for day and night data fusion and urban mapping.

1. Introduction

Urban areas, accounting for only 0.5% of the Earth's surface, accommodate more than 50% population (Schneider et al., 2009; UN, 2018). Abundant natural landscapes have been transformed into artificial surfaces through intensive human activities (Song et al., 2018b). Goal 11 “sustainable cities and communities”, i.e., “to make cities and human settlements inclusive, safe, resilient, and sustainable” was proposed in the Sustainable Development Goal (Klopp and Petretta, 2017). Under this background, these human activities as well as the related consequences should be monitored in time, and urban functional zone

(UFZ) can be considered as a suitable monitoring unit. UFZs are basic segmented regions carrying different socioeconomic activities, e.g., residence, commerce, and industry (Song et al., 2018a; Zhang et al., 2018). The region (or a city block) can be conveniently obtained based on the segmentation of road networks (Hu et al., 2016; Voltersen et al., 2014; Zhang et al., 2017). UFZ maps are fundamental for investigating urban expansion (Kane et al., 2014; Zhou et al., 2016) and redevelopment (Kwon et al., 2014; Wen et al., 2019), and moreover, they can also be used for health survey of urban residents (Bratman et al., 2012; Sallis et al., 2016), understanding urban environments such as microclimate (Huang and Wang, 2019; Peng et al., 2016a; Zhang et al., 2016),

* Corresponding author at: School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, PR China.

E-mail address: zjjerica@whu.edu.cn (J. Li).

<https://doi.org/10.1016/j.isprsjprs.2021.03.019>

Received 6 July 2020; Received in revised form 20 March 2021; Accepted 22 March 2021

Available online 6 April 2021

0924-2716/© 2021 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

hydrology (Li et al., 2018), anthropogenic heat (Sun et al., 2018), etc. Overall, it can be stated that functional zone mapping provides the basis for urban research and investigation.

UFZ mapping using high spatial resolution (HSR) images has drawn considerable attention since they are able to describe detailed land cover information. Existing methods mostly adopt two-dimensional (2D) visual features contained in images, e.g., spectral, textural, and geometrical attributes, to interpret functional zones (Hu and Wang, 2013; Song et al., 2018a; Wu et al., 2009). For instance, Zhang et al. (2015) used spectrum, gray level co-occurrence matrix, and shape features to classify urban functional zones in Beijing and Zhuhai. Liu et al. (2017a) adopted landscape metrics that quantify composition and configuration of the land covers (e.g., buildings and vegetation) to identify urban villages in Wuhan and Shenzhen. Meanwhile, the models that characterize the relationship between visual features were also adopted for UFZ mapping, e.g., bag-of-visual-words (BOVW). BOVW model views each functional zone as a bag of “visual words” with each “word” representing a typical visual pattern (Huang et al., 2018c; Peng et al., 2016b). By encoding the visual features with a series of “visual words”, functional zones can be effectively represented based on BOVW model. This model has been successfully applied to UFZ mapping in a number of Chinese cities (Liu et al., 2017b; Lu et al., 2018; Zhu et al., 2018). Besides BOVW, Zhang et al. (2017) proposed a three-level Bayesian model with each level characterizing visual features, object categories, object patterns as well as their hierarchical relations for recognizing urban functional zones. More recently, convolutional neural networks which can extract high abstract spatial features from images have been also utilized for UFZ mapping in several studies (Huang et al., 2018a; Zhou et al., 2020).

Summarizing the current literature, it can be found that most existing functional zone mapping methods focus on 2D (or planar) visual features. In fact, in addition to the 2D urban landscapes, urban structures also exhibit apparent three-dimensional (3D) characteristics, e.g., high-rise central business districts and low-rise shanty towns (Chen et al., 2014; Taubenböck et al., 2013). However, the potential of 3D characteristics in distinguishing urban functional zones has been less examined. Therefore, aided by the Ziyuan3-01 (ZY3-01) satellite, which can simultaneously acquire multi-view images (nadir, forward, and backward modes), this research aims to perform UFZ mapping by capturing the urban vertical information from the multi-view images. Compared to LiDAR (light detection and ranging) data, multi-view images have a variety of advantages in describing the urban 3D structures, e.g., low cost, rapid acquisition, wide coverage, and convenience for time-series updating (Huang et al., 2018b, 2017; Liu et al., 2019). In this context, one of the contributions of this study is to investigate the effect of multi-view satellite images for UFZ mapping.

On the other hand, nighttime light images are capable of reflecting human activity and socio-economic attributes (Elvidge et al., 2017; Levin and Duke, 2012; Levin et al., 2020; Zhang and Seto, 2011). Currently, the coarse or medium spatial resolution nighttime light imagery, e.g., the Defense Meteorological Satellite Program’s Operational Line-scan System (DMSP-OLS) (~1 km), Visible Infrared Imaging Radiometer Suite (VIIRS) (~500 m), and Luojia-1 satellite (~130 m), has been widely employed in global- or regional-scale applications, such as urban area extraction (Cao et al., 2009; Li and Zhou, 2017; Ou et al., 2019; Xie and Weng, 2016), socio-economic parameter estimation (Bennett and Smith, 2017; Mellander et al., 2015; Song et al., 2019), urban center detection (Chen et al., 2017; Zhang et al., 2013). However, it should be noted that these medium- to coarse-resolution images are not fine enough to exhibit the detailed structures of nighttime lights within a city, which, therefore, restricts their effectiveness for UFZ mapping. In recent years, the HSR nighttime light data become available, nevertheless, their relevant applications are limited to several studies concerning spatial pattern analysis of lights (Hale et al., 2013; Kuechly et al., 2012; Xue et al., 2020), identification of lighting source types (Zheng et al., 2018), and impervious area extraction (Tang et al.,

2020). In this research, therefore, we aim to investigate the effectiveness of the HSR nighttime light imagery (spatial resolution of 0.92 m) acquired from Jilin1-07 (JL1-07) satellite for the UFZ mapping. Moreover, we proposed a building enhanced nighttime light index (BENI) to sufficiently describe the characteristics of different urban functional zones from the HSR nighttime light imagery. To the best of our knowledge, few studies have considered HSR nighttime light images in the UFZ mapping.

In summary, we propose a daytime and nighttime data fusion method, i.e., fusion of HSR multi-view optical (ZY3-01) and nighttime light (JL1-07) imagery, for UFZ mapping, and these two data sources can depict the physical (2D and 3D urban structures) and socio-economic (fine-scale nighttime human activities) characteristics of functional zones, respectively. In this research, according to the characteristics of the study areas, six UFZ categories are considered: residential zones, shanty towns, parks and green space, industrial zones, commercial zones, and under-construction zones. Our research is able to provide new perspectives for urban functional zone mapping and HSR multi-source data fusion.

2. Study area and datasets

2.1. Study area

Two megacities of China, Beijing and Wuhan, were selected as the study areas for the UFZ mapping (Fig. 1). Beijing is the capital and the economic, political, and cultural center of China. In the context of the rapid urban expansion in Beijing, its urban area is surrounded by a series of ring roads around the Forbidden City. In this study, the urban core district within the 4th ring road, covering an area of ~300 km², was selected as our study area. The other city, Wuhan, is the capital of Hubei province and the industrial, education, and transportation center in Central China. Similarly, we selected the main urban area of Wuhan (~517 km²) as the other study area.

2.2. Datasets

The daytime optical and nighttime light images used in this study were acquired by ZY3-01 and JL1-07 satellites, respectively. The specific parameters of these two satellites are presented in Table 1.

The ZY3-01, the HSR three-line array stereo mapping satellite, carries one multi-spectral camera and three multi-view panchromatic cameras viewing towards the 22° backward, nadir, and 22° forward directions, respectively. In this way, ZY3-01 satellite acquires the multi-view images simultaneously, which can reduce the difference in atmospheric conditions and solar illumination among the multi-view images and focus on the angular variations and vertical information (Liu et al., 2019). The cloud-free multi-spectral and multi-view images were used, and the images of Beijing and Wuhan were acquired on 1st January 2018 and 28th October 2018, respectively.

The JL1-07 HSR nighttime satellite carries two multispectral cameras with on-board calibration devices and can provide fine-scale nighttime light information. To cover the study area in Beijing and Wuhan, 14 and 12 scenes of JL1-07 images were used, respectively. Specifically, 6 and 8 image scenes were collected on 29th March 2018 and 14th May 2018, respectively, and were mosaicked to cover the study area of Beijing. 12 scenes of Wuhan images were all collected on 7th April 2018. In addition, the side-swing angles of all these images were smaller than 7°, which can avoid geometrical difference between them. The mosaicked images after pre-processing (see Section 3.1) are shown in Fig. 1 (c) and (e).

2.3. Definitions of urban functional zones

In this study, the basic mapping unit is blocks, which were segmented by road networks. The primary-, secondary- and tertiary-

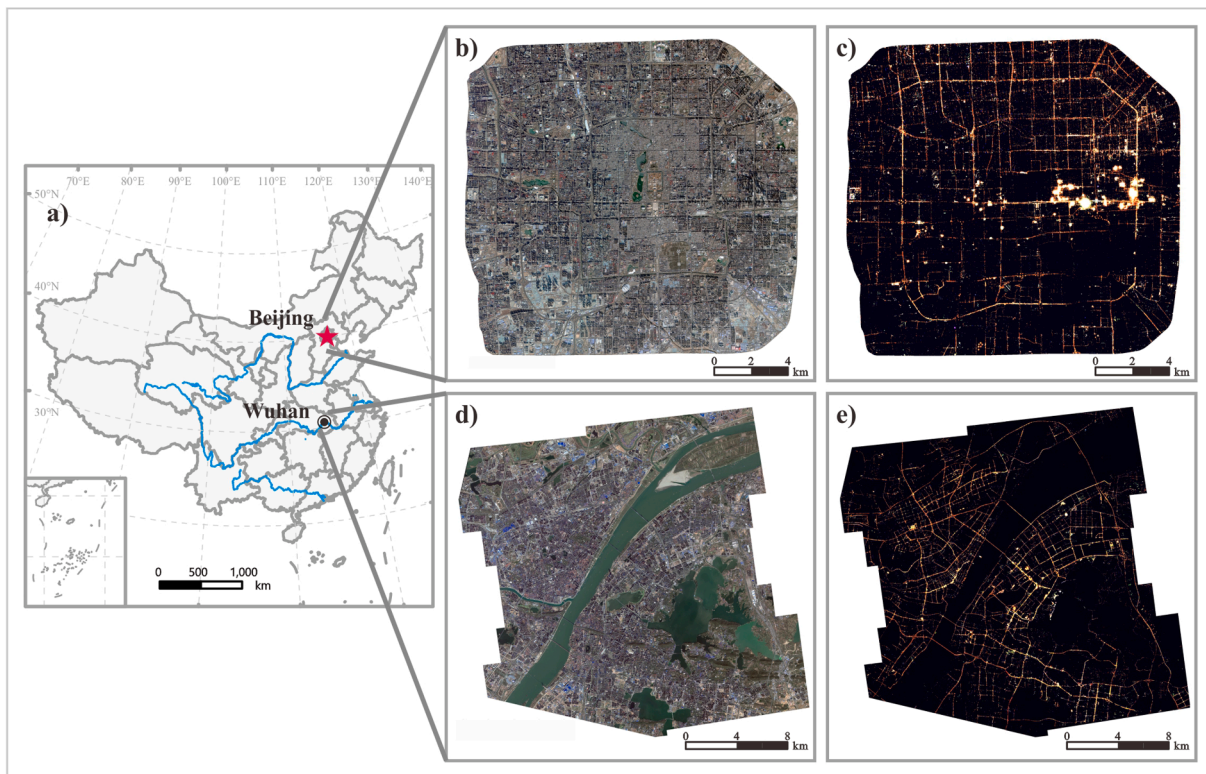


Fig. 1. Overview of the study areas and datasets: (a) presents the geographic locations of Beijing and Wuhan; (b) and (d) are their pansharpened ZY3-01 multi-spectral images; and (c) and (e) are nighttime light images acquired by JL1-07.

Table 1
Specific parameters of ZY3-01 and JL1-07 satellites. (PAN = panchromatic, MS = multi-spectral).

	ZY3-01	JL1-07
Launch time	9th January 2012	19th January 2018
Orbit altitude (km)	506	535
Orbit type	10:30 a.m.	10:30 a.m.
	Sun-synchronous	Sun-synchronous
Swath width (km)	51 × 51	11 × 4.5
Spatial resolution (m)	PAN-nadir: 2.1 PAN-forward: 3.5 PAN-backward: 3.5 MS: 5.8	MS: 0.92
Radiometric resolution (bits)	10	12
Wavelength (nm)	PAN: 500–800 Blue: 450–520 Green: 520–590 Red: 630–690 Infrared: 770–890	Camera1-Blue: 426–546 Camera1-Green: 494–598 Camera1-Red: 584–738 Camera2-Blue: 424–512 Camera2-Green: 490–588 Camera2-Red: 582–730

level road networks were obtained from OpenStreetMap (OSM) in August 2018. Since urban functions in the parcels segmented directly by OSM can be heterogeneous and mixed, e.g., residential zones and shanty towns in Case 1, and residential zones and parks in Case 2 (Fig. 2), these mixed zones were further separated into purer sub-blocks. In addition, several errors from the original OSM should be corrected manually, as demonstrated in Case 1. In this way, 1,537 and 1,717 UFZ blocks were generated in Beijing and Wuhan, respectively. In view of urbanization process and characteristics of both study areas, six categories of UFZs, including residential zones, shanty towns, parks and green space, industrial zones, commercial zones, and under-construction zones, were considered in this study. Their definitions are detailed in Table 2. The ground reference of UFZs was manually interpreted based on the field

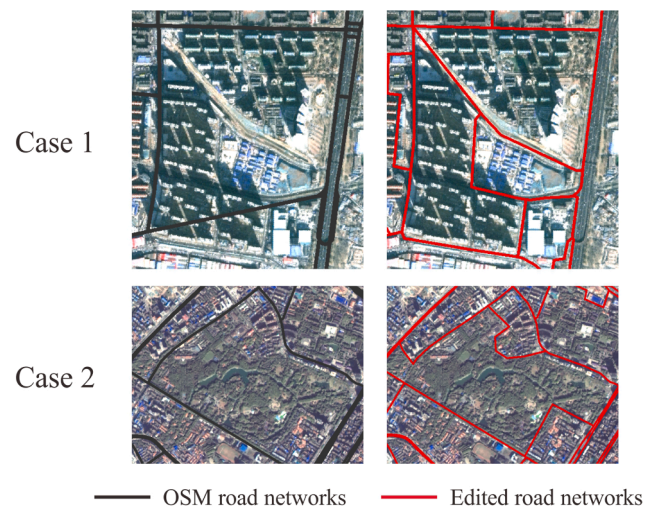


Fig. 2. Two representative examples indicating the manual editing of blocks.

Table 2
The definitions of the six types of functional zones in this study.

Types	Definitions
Residential zones	Well-planned housing area, such as houses and apartments
Shanty towns	Densely arranged temporary, old, or dilapidated buildings, such as urban villages
Parks and green space	Places for ecological and recreational purposes, such as parks or lawns
Industrial zones	Places to produce or store goods
Commercial zones	Places for business, trade, or sales, such as financial centers, shopping malls, and markets
Under-construction zones	Construction sites or vacant land

investigation and by referring to the open-source geographic information, e.g., points of interest and street-view images.

3. Methodology

In this study, a daytime and nighttime multi-source feature fusion method (Fig. 3) was proposed for UFZ mapping. Specifically, daytime features, i.e., spectral and multi-view feature, were extracted through the BOVW (bag of visual word) model and SDA (semi-supervised discriminant analysis) dimensionality reduction approach, and the nighttime feature consisted of nighttime light brightness values and the proposed building enhanced nighttime light index (BENI). In particular, BENI was designed to improve separability among different UFZ categories. Finally, spectral, multi-view, and nighttime light features were integrated for functional zone mapping. Details of the proposed method are introduced as follows.

3.1. Pre-processing

First, all the images were reprojected to the Universal Transverse Mercator system. ZY3-01 and JL1-07 images were radiometrically calibrated to the top of the atmosphere reflectance and radiance, respectively. Afterwards, the multi-spectral, ZY3-01 multi-view off-nadir were co-registered to the ZY3-01 nadir-view image using the first-order polynomial transform with bilinear resampling method. The RMSE (root mean square error) of the co-registration is less than half a pixel. Subsequently, the multi-spectral images were pan-sharpened by the Gram Schmidt transformation. The forward- and backward-view images were normalized to the nadir-view image by histogram matching. The nighttime light images of the two study areas were mosaicked separately. In specific, nighttime light images were co-registered to the nadir daytime optical images based on the matched point pairs (e.g., road junctions) with RMSE error less than one pixel, and the co-adjacent nighttime light images were mosaicked along the seam line (Li et al., 2019c). Finally, the multi-spectral, multi-view and nighttime light images were resampled to 2 m.

3.2. Multi-spectral and multi-view features

Spectral features were derived from the pan-sharpened multi-spectral bands, and multi-view features were originated from the forward-,

nadir-, and backward-view panchromatic bands. It should be noted that spatial displacement of ground objects in multi-view images is relevant to the height of objects and satellite viewing angles. As seen in Fig. 4, high-rise buildings show large differences among the multi-view images, e.g., the spatially displaced roofs and façade, but on the other hand, the low-rise architectures or ground objects exhibit small differences. Therefore, multi-view images are potential in delineating urban 3D structures. The spectral and multi-view bands were further processed by the BOVW models as shown in Fig. 5. In this study, each block was partitioned into a set of patches. “Visual words” were learned from these patches and each word represented a local feature or pattern. In this way, the multi-spectral and multi-view features for each functional zone were represented by the occurrence frequency of visual words that the patches were assigned to, and spectral and 3D visual information provided by the multi-spectral and multi-view images were encoded in the frequency histograms. The steps are detailed as follows:

- (1) Patch generation. Each block consists of a set of $N \times N$ m² image patches, which are mutually overlapped to some degree (overlapping of M meter). In this study, N was set to 32 m in consideration of the characteristics of the study areas and the spatial resolution of the images. Meanwhile, to achieve better feature representation of functional zones, the overlapping between adjacent patches was fixed to 24 m ($M = 24$). When the overlapping area between a patch and the block exceeds 80% of the patch area, this patch was kept as a part of the block, otherwise it would be excluded.
- (2) Patch-level feature descriptions. The spectral and multi-view features for each patch were generated through the histograms of 4 multi-spectral bands and 3 multi-view images, respectively. Each image band (either spectral or multi-view band) was represented by a 32-bin histogram, leading to 128- (32×4) and 96-bin (32×3) spectral and multi-view feature histograms, respectively.
- (3) Generation of visual words. 200,000 patches were randomly selected from the study areas, and the visual words, i.e., clustering centers, were then generated by k-means. Here, visual words for the spectral and multi-view features were acquired separately and their numbers of visual words were both set to 2048.

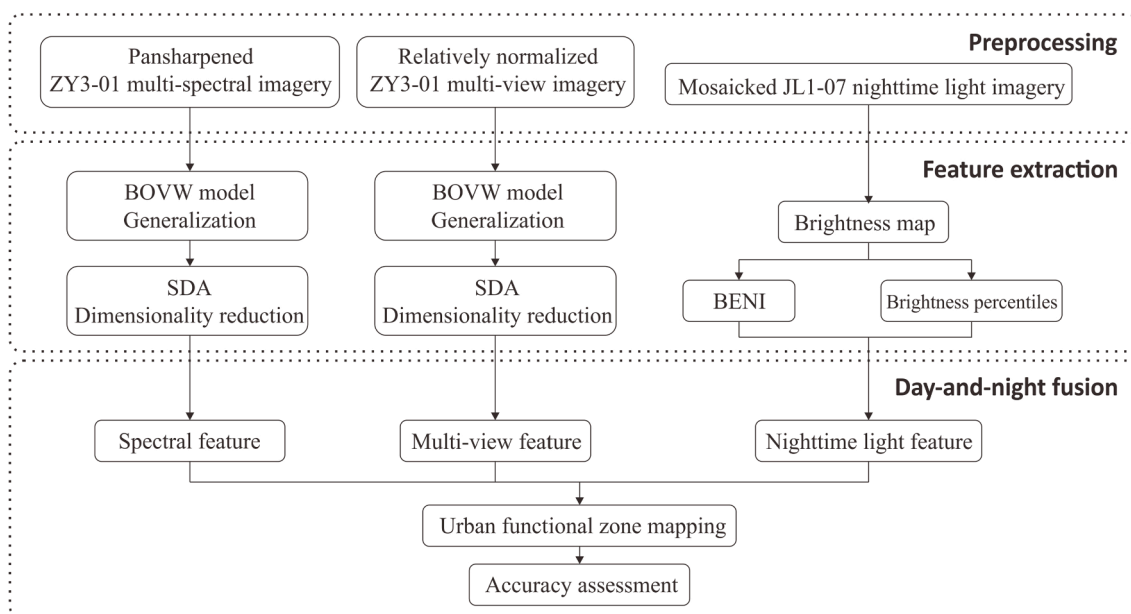


Fig. 3. Proposed workflow of the urban functional zone mapping.

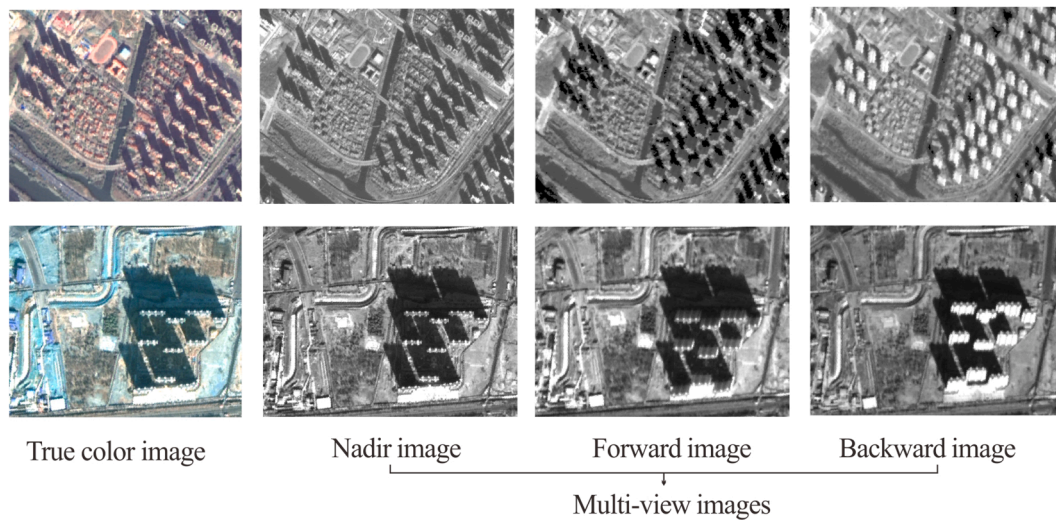


Fig. 4. Two examples of true color and multi-view images.

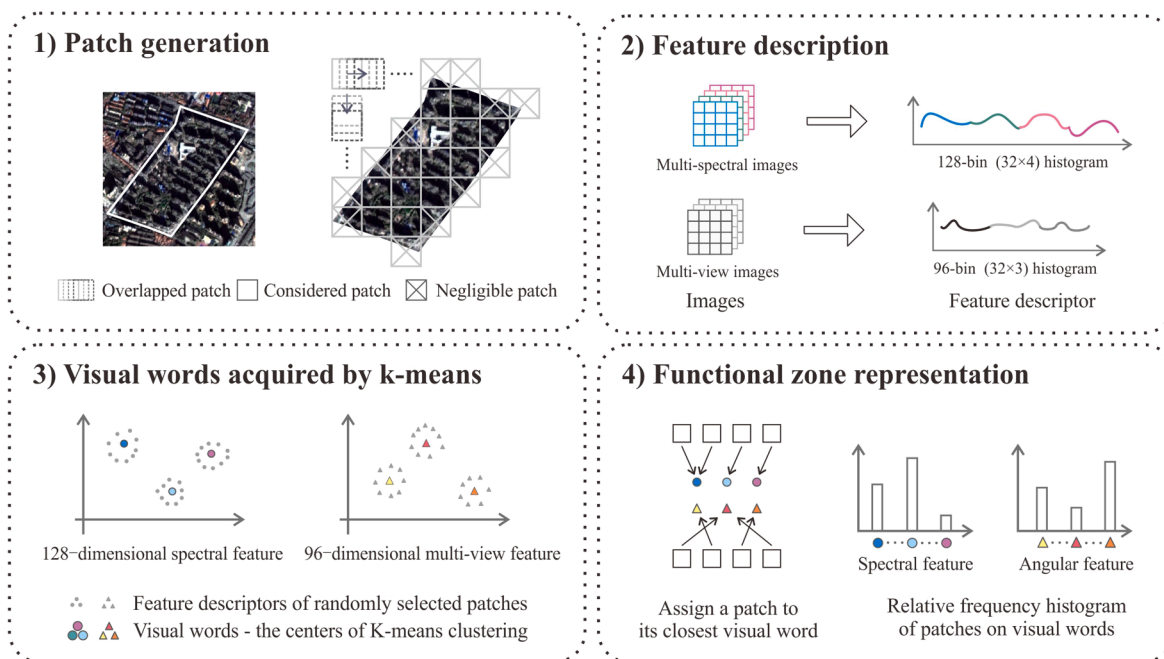


Fig. 5. Workflow of the spectral and multi-view feature extraction through the BOVW models.

- (4) UFZ feature representation. Each image patch was assigned to its nearest visual word in terms of the Euclidean distance in the feature space. Subsequently, each functional zone, i.e., a block that is composed of a set of overlapped patches, can be represented as the frequency histogram of the 2048 visual words in the block. Our UFZ feature extraction approach based on the BOVW models has two advantages. On one hand, the global features are effectively represented since visual words can capture a large number of typical patterns in the whole study area. On the other hand, the frequency histogram is appropriate for describing the features of various functional zones which have irregular shape and inconsistent sizes.
- (5) Dimensionality reduction. To obtain more representative spectral and multi-view features, semi-supervised discriminant analysis (SDA), a dimensionality reduction approach, was further adopted to refine the features generated by the BOVW models. SDA considers the features and the class labels of training set (see Section

3.5), which can reduce the feature dimensionality, and simultaneously preserve the separability between different classes (Cai et al., 2007). In this way, the original 2048-dimensional spectral and multi-view features that generated by the BOVW models were separately projected into a 256-dimensional subspace through the SDA algorithm.

3.3. Nighttime light features

Brightness is the most intuitive feature of nighttime light images, and in this study the brightness (BR) was calculated as (Zheng et al., 2018):

$$BR = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \quad (1)$$

where R , G , and B denote the radiance of red, green, and blue bands of JL1-07 images, respectively. Based on the brightness map, the nine deciles, i.e., 10th, 20th, ..., 80th and 90th percentiles, were calculated in each block to sufficiently describe the nighttime radiometric

information.

It should be noted that most of lights captured by the nighttime light sensors are from the upward light emission, but the indoor illuminations of buildings, cannot be effectively recorded (Katz and Levin, 2016; Zheng et al., 2018). Therefore, the brightness of some functional zones, such as residential zones and shanty towns, is usually underestimated in the nighttime light images, leading to underestimation of population density and human activity intensity in these zones. Therefore, we proposed a building enhanced nighttime light index (BENI) by synthesizing the nighttime light brightness and building density, in order to strengthen the brightness variations as well as the recognition for the functional zones with high building density but low brightness. BENI is calculated for each block and defined as:

$$BENI_i = \ln(b_{ri} + c) \times d_{Building_i} \quad (2)$$

$$b_{ri} = \frac{b_i}{\bar{b}} \quad (3)$$

$$\bar{b} = \frac{\sum_i^n b_i}{n} \quad (4)$$

where b_i is defined as the mean brightness of all pixels in block i and called as zone brightness thereafter, n is the total number of UFZs, \bar{b} is the average of zone brightness over all the UFZs, b_{ri} and $d_{Building_i}$ is relative brightness and building density for a certain block i , respectively, and c is a constant.

In the BENI, the building density ($d_{Building_i}$) was defined as the percentage of building area in the block i . Buildings were extracted automatically by the morphological building index (MBI). MBI considers the implicit relationship between the bright roof structure of buildings and the shadow around them, and it locally highlights building structures by representing the spectral-spatial characteristics with a series of morphological operators (Huang et al., 2014). MBI is able to directly extract building footprints from the multi-spectral images in an effective and automatic manner (Wen et al., 2016). MBI is defined as:

$$MBI = \frac{\sum_{s \in S} \sum_{d \in D} (DMP - WTH(s, d))}{N_s \times N_D} \quad (5)$$

where $DMP - WTH(s, d)$ means the differential morphological profiles of the white top-hat, s and d mean the direction and scale of the structural element, and N_s and N_D is the total number of scales and directions, respectively. More details of building extraction procedures can be referred to (Huang and Zhang, 2011). The other component of BENI, relative brightness (b_{ri}), is calculated as the zone brightness (b_i) divided by its average (\bar{b}). The numerical ranges of relative brightness and building density are similar and comparable (~ 1), which balances their weights when integrated in BENI. Moreover, considering that the distribution of zone brightness is highly right-skewed, a natural logarithmic transformation was applied to the relative brightness b_{ri} . The value of the constant term (c) was set as the constant e (≈ 2.718) in this study, to

ensure that the value after the logarithmic transformation is still around 1.

3.4. Day-night information fusion

As aforementioned, we can derive the spectral and multi-view features from daytime multi-view imagery through the BOVW models, and on the other hand, we can extract nighttime features, including deciles of brightness and BENI, from HSR nighttime light imagery. In the experiments, we would like to test the performance of daytime-nighttime data fusion for functional zone mapping with different feature sets as well as their combinations (Fig. 6):

(1) In case 1 and 2, the spectral and multi-view features were fed into the support vector machine (SVM) classifier, respectively.

(2) Case 3: Spectral-multi-view feature fusion, achieved by integrating the results of case 1 and 2 through the soft-voting strategy. Specifically, in the soft voting, the prediction probabilities for each class were averaged in a set of classifiers, and the class with the largest averaged probability was considered as the final prediction:

$$C(UFZ) = \underset{k=\{1, \dots, K\}}{\operatorname{argmax}} \left\{ \frac{1}{F} \sum_{f=1}^F p_f^k(UFZ_i) \right\} \quad (6)$$

where $C(UFZ_i)$ is the class label of functional zone i derived by soft voting, $k = 1, \dots, K$ refers to the class label output, F denotes the number of classifiers ($F = 2$ in this study), and $p_f^k(UFZ_i)$ is the probability output of classifier f for class k .

(3) Day-night data fusion, including spectral and nighttime light fusion (case 4), multi-view and nighttime light fusion (case 5), and combination of spectral, multi-view, and nighttime light features (case 6). Notably, in the case of 4 and 5, the multi-feature fusion was achieved by stacking (Huang and Zhang, 2010), while in case 6, the soft voting was utilized to fuse the classification results of case 4 and 5.

3.5. Accuracy assessment

The basic unit for accuracy assessment is a block. The block-based reference samples were manually interpreted in terms of field investigation or geographic information data, e.g., points of interest and street-view images. The reference samples were randomly divided into training and test set, respectively, as shown in Table 3. Moreover, all the classifications were repeated 40 times by randomly choosing different training sample sets, and their performances were evaluated by mean and standard deviation of overall accuracy (OA), class-specific accuracy, and confusion matrix.

4. Results

4.1. Mapping results of daytime features

Classification accuracies using different feature combinations are

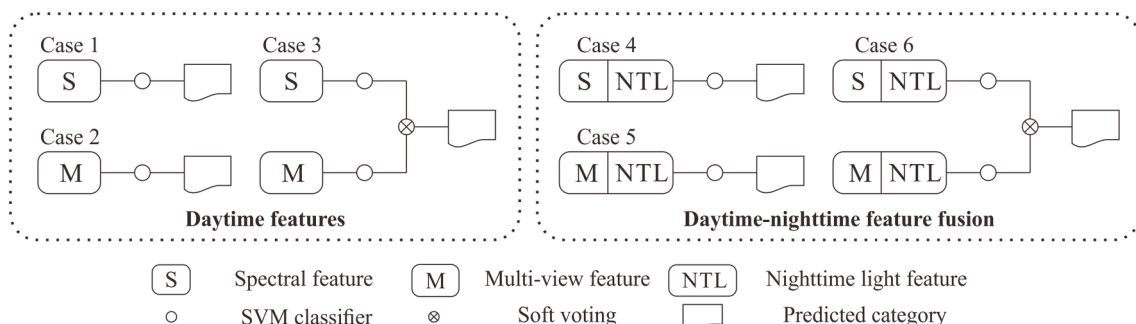


Fig. 6. Day-night data fusion for UFZ mapping with different features as well as their combinations.

Table 3
Number of training and test samples.

Categories	Beijing		Wuhan	
	Training	Test	Training	Test
Residential zones	35	752	45	706
Shanty towns	35	109	40	93
Parks and green space	35	90	40	226
Industrial zones	30	69	40	118
Commercial zones	30	249	35	133
Under-construction zones	35	68	40	201
Total	200	1337	240	1477

shown in Table 4 (Beijing) and Table 5 (Wuhan) and the mapping results are presented in Fig. 7. In general, OA achieved by the individual use of spectral or multi-view features was around 80% in both Beijing and Wuhan, but the performances of these two features varied slightly in the two study areas. Specifically, in Wuhan, spectral features outperformed multi-view features for all the UFZ categories. In Beijing, however, spectral features obtained better accuracies for parks and green space, and under-construction zones, while multi-view features were better for residential and commercial zones.

Results showed that the OA of both study areas was raised, when combining spectral and multi-view features. The increment of OA was 5.3% and 1.9% in Beijing, compared to the individual use of spectral and multi-view features, respectively, and the corresponding increment of OA was 2.2% and 6.3% in Wuhan. By fusing the two features, the class-specific accuracy of residential zones, shanty towns, and industrial zones in Beijing was increased by 6.5%, 3.1%, and 5.8%, respectively, as compared to that of using only spectral features; but for the categories of parks and green space, and under-construction zones, their increments were slight, (0.1% and 0.6%, respectively). With regard to Wuhan, all the categories except for commercial zones had an obvious increase by 1.5% ~ 4.5%. These results revealed that: 1) the multi-view features extracted from the multi-view images were effective in interpretation of functional zones, and 2) spectral-multi-view feature fusion was beneficial for improving the accuracy. Notably, in both Beijing and Wuhan, commercial zones corresponded to the lowest accuracy among all the UFZs, and the effectiveness of identifying this category was limited in spite of fusing spectral and multi-view features. This implied that commercial zones could not be well recognized using only daytime images.

Table 4
Classification accuracies (%) using different feature combinations in Beijing.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
	S	M	S + M	S + NTL	M + NTL	S + M + NTL
Res.	84.4 ± 5.2	90.4 ± 3.7	90.9 ± 3.7	91.1 ± 2.8	93.3 ± 1.9	93.8 ± 2.0
Sha.	87.5 ± 4.4	87.9 ± 4.1	90.6 ± 3.7	87.4 ± 3.7	88.4 ± 4.2	90.6 ± 3.3
Par.	89.4 ± 3.5	80.1 ± 5.2	89.5 ± 3.3	89.1 ± 3.4	82.0 ± 5.2	89.3 ± 3.9
Ind.	79.9 ± 5.0	81.8 ± 4.6	85.7 ± 4.6	81.1 ± 5.7	84.6 ± 4.1	86.5 ± 5.0
Com.	49.8 ± 6.2	56.1 ± 5.4	55.6 ± 5.5	78.1 ± 4.5	78.0 ± 4.4	79.4 ± 4.3
Und.	82.0 ± 6.1	69.7 ± 6.7	82.6 ± 5.0	82.6 ± 5.2	70.5 ± 6.1	82.6 ± 4.8
OA	78.2 ± 2.4	81.6 ± 1.8	83.5 ± 1.5	87.3 ± 1.5	87.7 ± 1.1	89.6 ± 1.1
Kappa	66.7 ± 2.9	71.6 ± 2.4	74.5 ± 2.0	80.4 ± 2.2	80.9 ± 1.6	83.8 ± 1.6

Note: S – spectral features, M – multi-view features, NTL – nighttime light features
Numbers in bold indicate the highest accuracy in each line.

Table 5
Classification accuracies (%) using different feature combinations in Wuhan.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
	S	M	S + M	S + NTL	M + NTL	S + M + NTL
Res.	87.1 ± 2.3	83.6 ± 4.1	89.5 ± 2.3	87.9 ± 2.3	86.0 ± 3.6	90.5 ± 2.2
Sha.	84.9 ± 5.3	81.9 ± 5.3	87.8 ± 4.6	85.2 ± 5.3	82.4 ± 5.4	87.9 ± 4.3
Par.	85.3 ± 3.9	82.9 ± 4.5	86.8 ± 4.0	85.6 ± 3.6	83.8 ± 4.3	87.5 ± 3.8
Ind.	77.5 ± 4.7	75.7 ± 5.5	81.9 ± 5.4	78.3 ± 4.7	75.8 ± 5.6	82.6 ± 5.4
Com.	54.3 ± 5.5	41.8 ± 6.2	53.7 ± 5.4	62.2 ± 5.4	54.2 ± 5.3	63.9 ± 3.9
Und.	76.3 ± 4.6	72.4 ± 3.5	78.5 ± 3.5	76.9 ± 4.7	72.7 ± 3.6	78.6 ± 3.6
OA	81.5 ± 1.1	77.4 ± 1.7	83.7 ± 1.1	82.8 ± 1.0	79.9 ± 1.6	85.2 ± 1.0
Kappa	74.5 ± 1.4	69.0 ± 2.0	77.3 ± 1.4	76.3 ± 1.3	72.4 ± 2.0	79.5 ± 1.3

Note: S – spectral features, M – multi-view features, NTL – nighttime light features

Numbers in bold indicate the highest accuracy in each line.

4.2. Mapping results of day-night feature fusion

As shown in Table 6 and 7, compared to the daytime feature fusion (case 3), the increment of OA achieved by the spectral-multi-view-nighttime feature fusion (case 6) was 6.1% and 1.5%, respectively, in Beijing and Wuhan. The day-night fusion gave the highest OA: 89.6% in Beijing and 85.2% in Wuhan.

In detail, the addition of nighttime light features significantly improved the accuracy of commercial zones by 23.8% in Beijing and 10.2% in Wuhan. Referring to the confusion matrices in Table 6 and 7, it can be seen that the addition of nighttime light features could greatly reduce the confusion between commercial and residential (or industrial) zones. For example, the commercial zone at the bottom left of the example P1 (Fig. 8), which is composed of an arena, two shopping centers, and several office buildings, was misclassified into an industrial zone when only daytime features were used. However, it was correctly classified as a commercial zone using day-night feature fusion since this region exhibits abundant decorative lights of the shopping malls in the nighttime light imagery. Similarly, the two commercial zones in the right side of the example P3 (Fig. 8), are shopping and entertainment hotspots in Beijing, and both were misclassified as residential zones using only daytime features. But the addition of nighttime features could effectively correct this error. In addition, day-night feature fusion also improved the detection accuracy of residential zones, and the increments were 2.9% and 1.0% in Beijing and Wuhan, respectively. For instance, when only daytime features were considered, the high-rise residential zone (the example P5 in Fig. 8) was misclassified as a commercial zone, but it was correctly classified by the daytime-nighttime fusion.

5. Discussion

5.1. Influence of the patch size

Patch size is an important parameter of the BOVW model. A small size is not sufficient to describe the objects as well as their contextual information in a block, but a large one may ignore the detailed structures of the block. To discuss the influence of the patch sizes, the classification results with different sizes, e.g., 16 × 16 m², 32 × 32 m², and 64 × 64 m², were compared. As seen from Tables 8 and 9, the patch size of 32 × 32 m² achieved the highest accuracy in Beijing and the second-best accuracy in Wuhan. Please notice that a smaller patch size corresponds to more complicated representation of a block, and hence leads

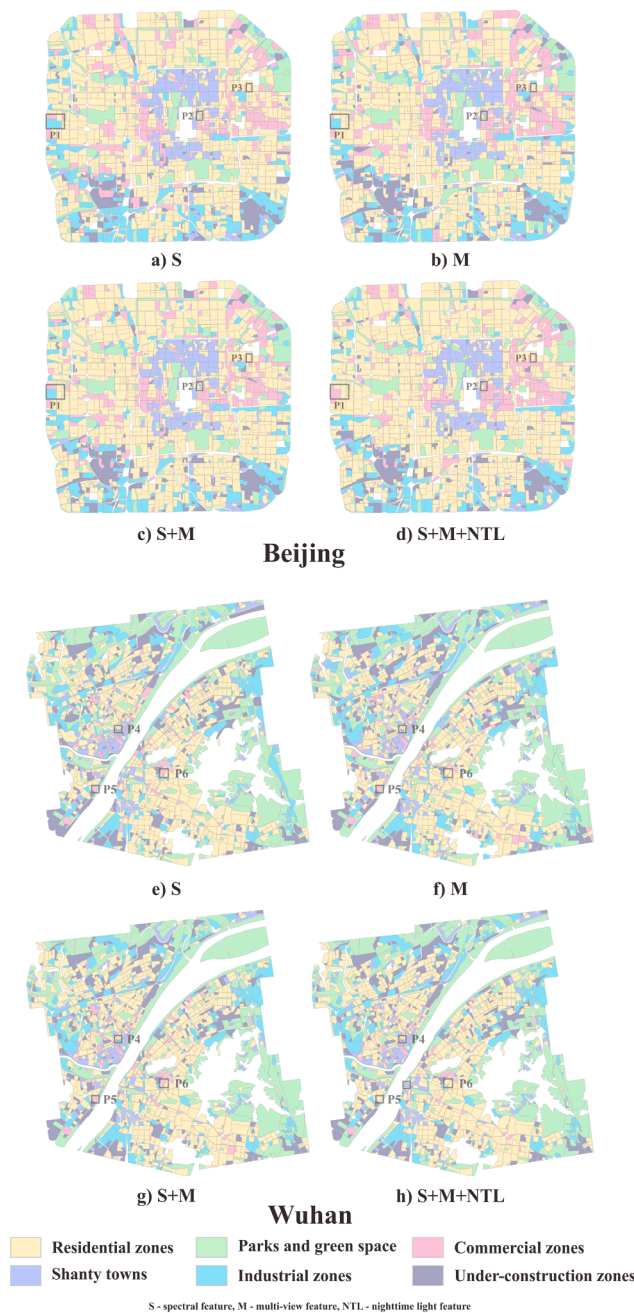


Fig. 7. Classification results of functional zones in Beijing (a-d) and Wuhan (e-h).

Table 6
Confusion matrix for functional zone mapping results in Beijing.

	Case 3: S + M					Case 6: S + M + NTL						
	Res.	Sha.	Par.	Ind.	Com.	Und.	Res.	Sha.	Par.	Ind.	Com.	Und.
Res.	684 ± 28	1 ± 1	1 ± 1	3 ± 1	48 ± 14	0 ± 1	705 ± 15	2 ± 1	1 ± 1	3 ± 2	20 ± 7	0 ± 1
Sha.	3 ± 2	99 ± 4	1 ± 1	1 ± 1	2 ± 1	0 ± 0	2 ± 2	99 ± 4	0 ± 1	1 ± 1	1 ± 1	0 ± 0
Par.	4 ± 3	1 ± 1	81 ± 3	1 ± 1	2 ± 1	4 ± 2	3 ± 2	1 ± 1	80 ± 3	1 ± 1	0 ± 1	4 ± 2
Ind.	10 ± 5	6 ± 4	2 ± 2	59 ± 3	54 ± 8	7 ± 3	10 ± 6	6 ± 4	1 ± 1	60 ± 3	27 ± 8	6 ± 3
Com.	51 ± 25	1 ± 1	1 ± 1	3 ± 2	138 ± 14	1 ± 1	30 ± 11	1 ± 1	2 ± 1	2 ± 1	198 ± 11	2 ± 1
Und.	2 ± 2	0 ± 1	5 ± 3	3 ± 2	5 ± 2	56 ± 3	2 ± 3	0 ± 1	5 ± 3	3 ± 2	4 ± 2	56 ± 3
	OA: 83.5 ± 1.6%					OA: 89.6 ± 1.1%						

Note: S – spectral features, M – multi-view features, NTL – nighttime light features.

to more computational burden. Therefore, considering the trade-off between accuracy and computational cost, the patch size of $32 \times 32 \text{ m}^2$ was chosen in our experiments.

5.2. Comparison of composite-brightness and nighttime multi-spectral bands

The nighttime features, including nighttime light brightness values and the proposed BENI, were extracted based on composite-brightness band (see Eq. (1)), but not nighttime multi-spectral bands. To investigate whether the nighttime multi-spectral bands can provide additional information, the results by using composite-brightness and multi-spectral bands are compared in Tables 10 and 11. As seen from the tables, the accuracy improvement brought by the nighttime multi-spectral bands is very marginal, only 0.2–0.5% in OA. This phenomenon can be attributed to the high correlation (0.91 and 0.99 at average in Beijing and Wuhan, respectively) among the nighttime multi-spectral bands. Therefore, it is appropriate to merge the nighttime multi-spectral bands to a pan-chromatic band in this research.

5.3. Contributions of daytime and nighttime features

In the proposed daytime-nighttime data fusion framework, the daytime and nighttime images can characterize the UFZ from different aspects. Spectral and multi-view features delineate the physical characteristics of UFZ, from the aspects of color and 3D structures, respectively. On the other hand, nighttime light features can be viewed as an effective complement by describing the nighttime human activities. In order to quantitatively evaluate the contributions of spectral, multi-view and nighttime light features in UFZ mapping, the permutation importance (Breiman, 2001) was used to measure their importance, as presented in Fig. 9.

In general, the spectral features made the largest contribution (49.3% in Beijing and 51.7% in Wuhan), followed by the multi-view (34.7% (Beijing) and 37.1% (Wuhan)) and nighttime light (16.0% (Beijing) and 11.2% (Wuhan)) features. For both study areas, the spectral and multi-view features had stable and important contributions to identification of all the UFZ categories except for the commercial zones, and the spectral attributes had a slightly higher weight than the multi-view features. This implied that the spectral and multi-view features from the daytime imagery were the most important information sources for UFZ mapping. However, nighttime light features showed the greatest contribution for recognizing the commercial zones, accounting for 65.3% and 43.4% in Beijing and Wuhan, respectively. Meanwhile, nighttime light features were also effective in identification of residential zones (Beijing), parks and green space (both cities), and industrial zones (Beijing). The contribution of nighttime light features in Beijing (16.0%) was much more than that in Wuhan (11.2%), and the increment of OA in Beijing (6.1%) brought by the nighttime light images was also larger than that in Wuhan (1.5%). The different performances of nighttime features in the two cities can be attributed to the proportion of lit pixels in the nighttime light images. In specific, pixels with minimum

Table 7
Confusion matrix for functional zone mapping results in Wuhan.

	Case 3: S + M						Case 6: S + M + NTL					
	Res.	Sha.	Par.	Ind.	Com.	Und.	Res.	Sha.	Par.	Ind.	Com.	Und.
Res.	632 ± 16	4 ± 2	3 ± 2	5 ± 3	24 ± 6	5 ± 3	639 ± 16	4 ± 2	2 ± 2	5 ± 3	17 ± 4	5 ± 3
Sha.	23 ± 9	82 ± 4	0 ± 0	2 ± 1	5 ± 2	1 ± 1	22 ± 9	82 ± 4	0 ± 0	2 ± 2	4 ± 2	1 ± 1
Par.	4 ± 3	1 ± 1	196 ± 9	3 ± 2	2 ± 1	11 ± 4	4 ± 3	1 ± 1	198 ± 9	4 ± 3	1 ± 1	10 ± 5
Ind.	21 ± 10	6 ± 3	11 ± 4	97 ± 6	24 ± 6	21 ± 5	21 ± 10	6 ± 3	11 ± 4	98 ± 6	21 ± 6	20 ± 5
Com.	17 ± 7	0 ± 0	0 ± 1	4 ± 2	71 ± 7	6 ± 3	10 ± 3	0 ± 0	0 ± 1	3 ± 2	85 ± 5	6 ± 3
Und.	9 ± 7	0 ± 1	15 ± 7	7 ± 3	7 ± 3	158 ± 7	9 ± 7	0 ± 1	15 ± 7	7 ± 3	6 ± 3	158 ± 7
	OA: 83.7 ± 1.1%						OA: 85.2 ± 1.0%					

Note: S – spectral features, M – multi-view features, NTL – nighttime light features

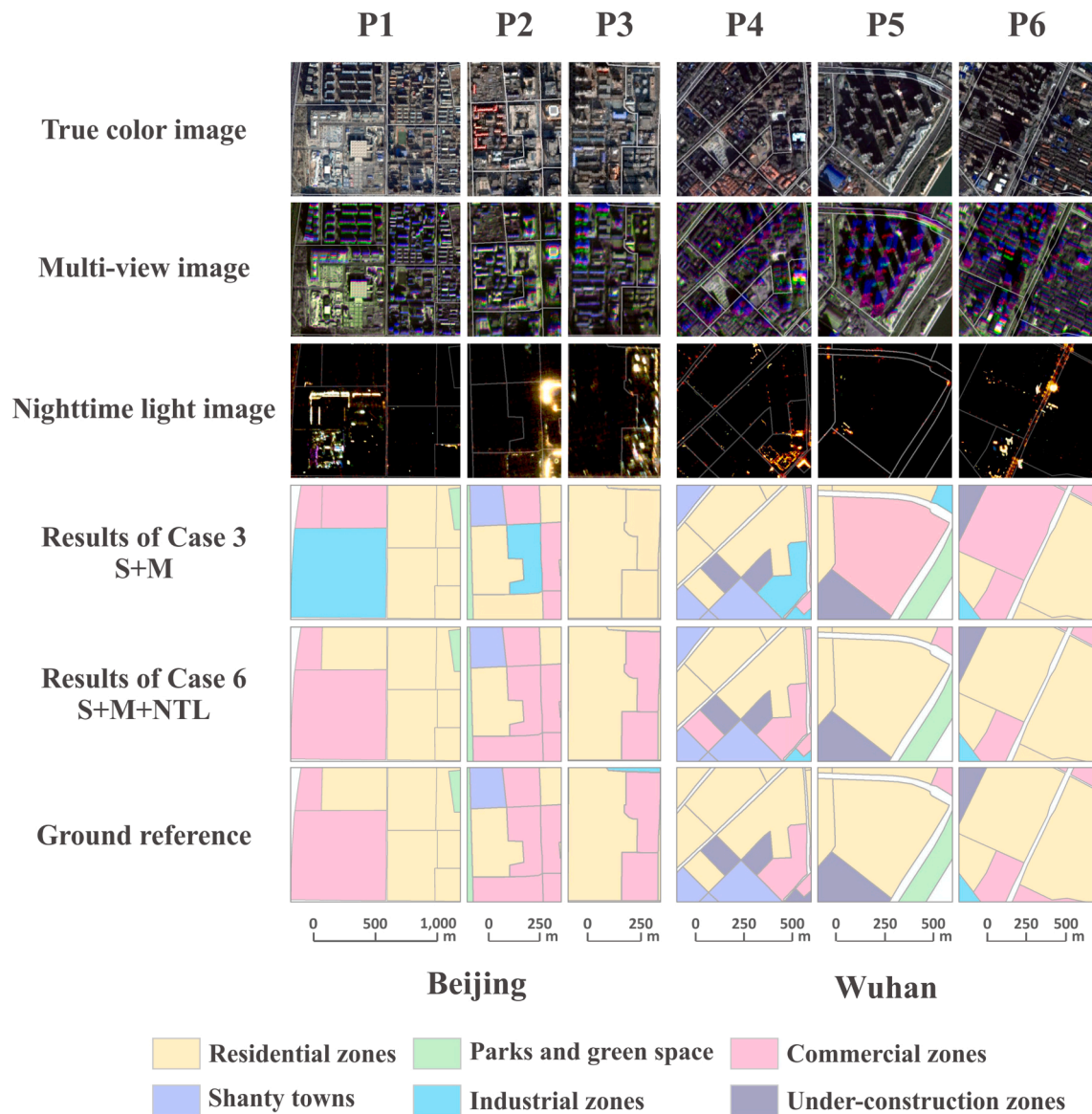


Fig. 8. Six representative examples (P1-P6) of classification results in Beijing and Wuhan. In the multi-view images, the nadir, forward, and backward panchromatic images are displayed as RGB composite. The locations of the six examples are marked in Fig. 7.

radiance can be regarded as the background, corresponding to the regions without or with extremely dim lights. In contrast, other pixels are considered as lit ones. The ratio of lit pixels can be regarded as an indicator to measure the information content of the nighttime light images, and a higher ratio signifies more information. In this study, the

ratio of lit pixels in Beijing and Wuhan was 6.8% and 5.0%, respectively, showing that the Beijing nighttime images are more informative. Furthermore, a larger ratio means smaller degree of skewness for the zone brightness and better discriminatory power provided by BENI. Accordingly, it can be said that the images with higher ratio of lit pixels

Table 8
Classification accuracies (%) of S + M + NTL using different patch sizes in Beijing.

Patch Size (m ²)	16 × 16	32 × 32	64 × 64
Res.	93.8 ± 1.9	93.8 ± 2.0	94.4 ± 1.9
Sha.	91.7 ± 3.0	90.6 ± 3.3	86.9 ± 4.1
Par.	89.9 ± 4.4	89.3 ± 3.9	89.2 ± 4.3
Ind.	83.8 ± 4.1	86.5 ± 5.0	85.9 ± 3.7
Com.	79.3 ± 4.4	79.4 ± 4.3	77.5 ± 5.3
Und.	82.6 ± 6.2	82.6 ± 4.8	80.1 ± 6.4
OA	89.6 ± 1.0	89.6 ± 1.1	89.1 ± 1.0
Kappa	83.7 ± 1.5	83.8 ± 1.6	83.0 ± 1.5

Note: S – spectral features, M – multi-view features, NTL – nighttime light features.

Table 9
Classification accuracies (%) of S + M + NTL using different patch sizes in Wuhan.

Patch Size (m ²)	16 × 16	32 × 32	64 × 64
Res.	91.2 ± 1.7	90.5 ± 2.2	89.5 ± 2.6
Sha.	91.2 ± 3.9	87.9 ± 4.3	81.4 ± 5.4
Par.	85.4 ± 3.8	87.5 ± 3.8	88.3 ± 4.0
Ind.	80.7 ± 5.1	82.6 ± 5.4	79.9 ± 4.8
Com.	68.2 ± 4.7	63.9 ± 3.9	59.3 ± 6.7
Und.	78.7 ± 3.7	78.6 ± 3.6	74.6 ± 4.4
OA	85.7 ± 0.9	85.2 ± 1.0	83.3 ± 1.2
Kappa	80.2 ± 1.2	79.5 ± 1.3	76.9 ± 1.5

Note: S – spectral features, M – multi-view features, NTL – nighttime light features.

Table 10
Classification accuracies (%) of S + M + NTL using nighttime composite-brightness and multi-spectral bands in Beijing.

	Composite-brightness	Multi-spectral bands
Res.	93.8 ± 2.0	94.2 ± 1.7
Sha.	90.6 ± 3.3	90.4 ± 3.5
Par.	89.3 ± 3.9	90.3 ± 3.6
Ind.	86.5 ± 5.0	86.4 ± 4.8
Com.	79.4 ± 4.3	79.4 ± 3.9
Und.	82.6 ± 4.8	82.6 ± 4.9
OA	89.6 ± 1.1	89.9 ± 1.1
Kappa	83.8 ± 1.6	84.2 ± 1.6

Note: S – spectral features, M – multi-view features, NTL – nighttime light features.

Table 11
Classification accuracies (%) of S + M + NTL using nighttime composite-brightness and multi-spectral bands in Wuhan.

	Composite-brightness	Multi-spectral bands
Res.	90.5 ± 2.2	90.8 ± 2.2
Sha.	87.9 ± 4.3	87.7 ± 4.1
Par.	87.5 ± 3.8	88.0 ± 3.6
Ind.	82.6 ± 5.4	82.9 ± 5.4
Com.	63.9 ± 3.9	66.9 ± 4.1
Und.	78.6 ± 3.6	78.4 ± 3.4
OA	85.2 ± 1.0	85.7 ± 1.0
Kappa	79.5 ± 1.3	80.2 ± 1.3

Note: S – spectral features, M – multi-view features, NTL – nighttime light features

are more effective in UFZ mapping, which can partly explain the different performances of nighttime features in Beijing and Wuhan.

Notably, the ratio of lit pixels is related to a series of factors, e.g., spatial configuration of lighting facilities as well as their mode (on or off), occlusion by other objects, atmospheric transmission of lights, and observation angles (Aubé, 2015; Katz and Levin, 2016). It is meaningful

to further investigate the difference of nighttime images between different cities in future work.

5.4. Efficacy and potentials of BENI

As introduced in Section 3.3, BENI was designed to mitigate the underestimation of nighttime light in the images, especially in the residential areas. The effectiveness of BENI is further analyzed here, in comparison to the original zone brightness. First, we calculated their skewness. Skewness is a statistical measure of how much the data distribution is asymmetrical from the normal distribution. A positive skewness means the most values are distributed on the right of the mean, and larger value corresponds to high skewness of the data distribution (Groeneveld and Meeden, 1984). The skewness of zone brightness and BENI is 10.64 and 1.50 in Beijing and 9.47 and 2.04 in Wuhan. Obviously, BENI greatly mitigates the skew of zone brightness, which shows that BENI can enlarge the variation of nighttime light brightness.

To show the superiority of BENI, we illustrate their distribution (Fig. 10), where all the reference samples (both training and test set) were involved in the statistics. It can be seen that compared to the original zone brightness, the proposed BENI can preserve good separability between commercial zones and other UFZs, and meanwhile enhance the contrast between low-brightness functional zones.

Moreover, to validate the advantage of the proposed BENI, additional experiments were conducted. The zone brightness and BENI were separately used under different scenarios (Table 12), and the paired *t*-test was employed to determine whether BENI significantly outperforms the zone brightness. As indicated by Table 12, in all the scenarios, BENI achieved higher OA than the zone brightness, and the increments in Beijing and Wuhan were 0.8–1.2% and 0.7–1.5%, respectively. All these accuracy increments were statistically significant according to the *t*-test.

The good performances of BENI have been confirmed in the experiments, but it is still potential in other aspects. For instance, BENI can be applied to other nighttime light images. During the calculation of BENI, the natural logarithmic transformation can mitigate the highly right-skewed effect of brightness, and the normalization operation further ensures its applicability. Therefore, BENI can be useful in other high-resolution nighttime light images that also are skewed in data distribution, such as aerial nighttime light images (Kuechly et al., 2012) and Luojia-1 satellite imagery (Li et al., 2019b), since it is able to correct the underestimation of lights in regions with high building density but low light brightness (e.g., residential zones). In addition, BENI can be considered as a data or image pre-processing method to correct the skewness of nighttime light brightness, which is beneficial for subsequent applications.

5.5. Limitations and future works

As discussed above, this study achieved accurate interpretation of functional zones through the proposed daytime-nighttime data fusion method. However, several limitations remain and deserve future research. For example, the detailed spatial pattern of lights can be further explored in high-resolution nighttime light features. Our study focused on brightness of nighttime lights, but their spatial patterns, such as shape, aggregation, and distribution, are also important descriptors of nightscape. Notably, the deep learning technology (Li et al., 2019a) can be used to mine HSR spectral, multi-view, and light information for a deep fusion. However, considering that the HSR nighttime light images have been rarely used in existing studies, this research aimed at analyzing the characteristics and performances of the relatively original features without the learning and optimization by the deep learning. In the future, we plan to investigate deep fusion of multi-source daytime and nighttime images. Last but not least, the application of the proposed method in other cities will be considered as future work.

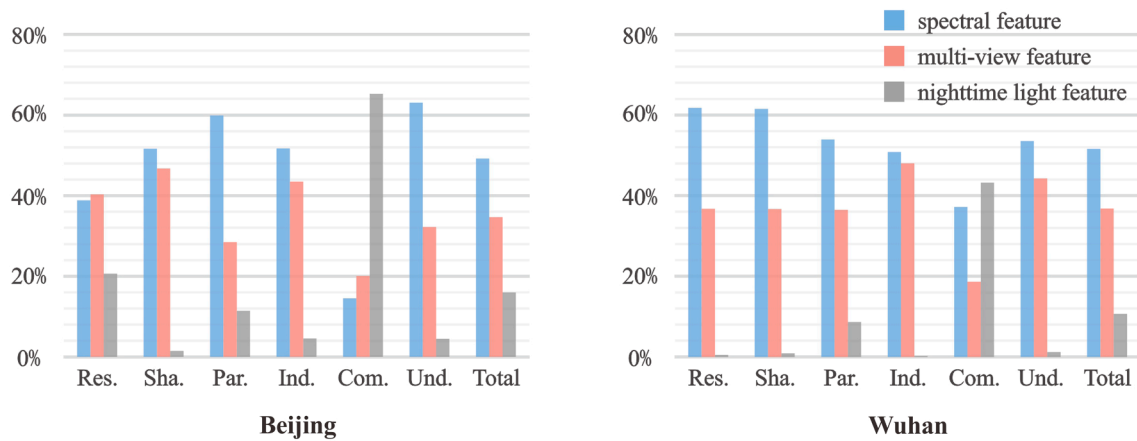


Fig. 9. The contributions of spectral, multi-view, and nighttime light features for functional zone mapping in Beijing and Wuhan.

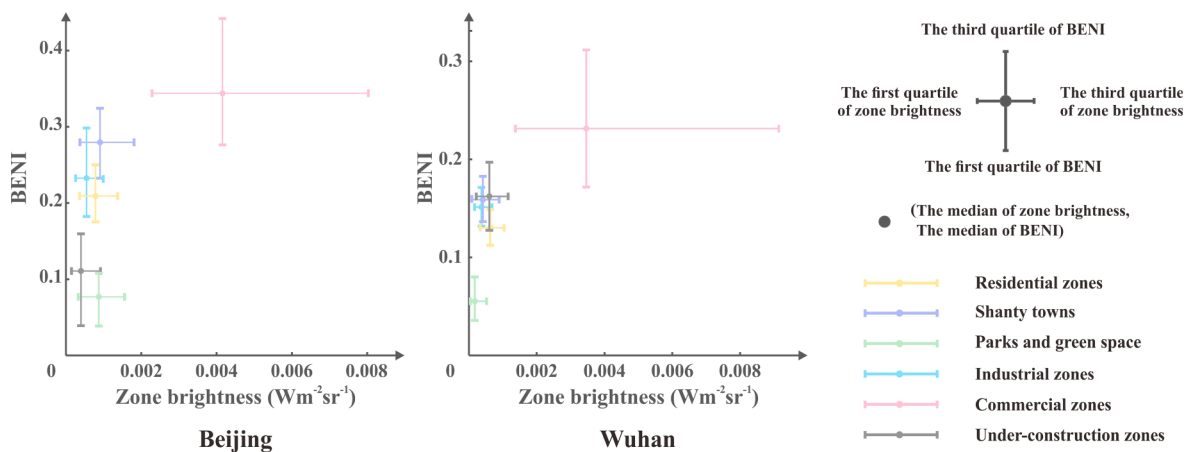


Fig. 10. Distribution of zone brightness and BENI for different functional zone categories.

Table 12

The comparison of overall accuracy (%) using zone brightness and BENI in different classification scenarios (S: spectral features, M: multi-view features, NTL: zone brightness or BENI).

	Beijing			Wuhan		
	NTL = zone brightness	NTL = BENI	Increment	NTL = zone brightness	NTL = BENI	Increment
S + NTL	83.1 ± 1.8	83.9 ± 1.8	0.8**	82.0 ± 1.0	82.7 ± 1.0	0.7**
M + NTL	84.6 ± 1.6	85.8 ± 1.6	1.2**	78.3 ± 1.7	79.8 ± 1.6	1.5**
S + M + NTL	87.0 ± 1.3	87.8 ± 1.3	0.8**	84.2 ± 1.1	85.0 ± 1.0	0.8**

** indicates significance at 0.01 confidence level using paired t-test.

6. Conclusions

This study proposed a daytime-nighttime data fusion for urban functional zone (UFZ) mapping. In this fusion, ZY3-01 multi-view imagery provided daytime characteristics including color and 2D/3D urban structures, while JL1-07 nighttime light images described high-resolution nighttime human activities. To our knowledge, this is the first time that the fusion of HSR nighttime light imagery as well as daytime imagery is used in urban classification and mapping. Meanwhile, HSR multi-view images can effectively delineate urban 3D structures, but they have been rarely used in UFZ mapping. Moreover, we proposed a building enhanced nighttime light index (BENI), which was designed specifically to overcome the underestimation of light intensity in the regions with high building density but low brightness.

Based on the experiments in the two megacities of China, Beijing and

Wuhan, the effectiveness of the proposed daytime-nighttime data fusion and the BENI has been validated. Our experimental results showed that the addition of multi-view features effectively improved the overall accuracy (OA) of UFZ mapping using the spectral features alone, and the increments were 5.3% and 2.2% in Beijing and Wuhan, respectively. Furthermore, when nighttime light features were added into the daytime features, OA was improved by 6.1% and 1.5% in Beijing and Wuhan, respectively. In particular, the use of nighttime light features achieved significant increase of accuracy for the commercial zones by 23.8% in Beijing and 10.2% in Wuhan, compared to only using the daytime features. The misclassifications between commercial and residential/industrial zones were substantially reduced.

In addition, the importance of daytime and nighttime features in UFZ mapping was analyzed. The contribution degrees of spectral and multi-view features were 49.3% and 34.7% in Beijing and 51.7% and 37.1% in

Wuhan, respectively. The contribution of nighttime light features was 16.0% in Beijing and 11.2% in Wuhan, which can effectively complement the daytime features. Moreover, better performances of the proposed BENI were observed compared to the original brightness, showing that BENI was able to correct the skewness of the high-resolution nighttime light data, and enhance the separability between various function zones. In general, our research provides a new and interesting paradigm, i.e., daytime-nighttime data fusion, for UFZ classification, and also provides new insights into the use of high-resolution nighttime light imagery.

Declaration of Competing Interest

The author declare that there is no conflict of interest.

Acknowledgements

The research was supported by the National Natural Science Foundation of China under Grants 42071311, 41771360 and 41971295.

References

- Aubé, M., 2015. Physical behavior of anthropogenic light propagation into the nocturnal environment. *Philos. Trans. R. Soc. B Biol. Sci.* 370 <https://doi.org/10.1098/rstb.2014.0117>.
- Bennett, M.M., Smith, L.C., 2017. Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics. *Remote Sens. Environ.* 192, 176–197. <https://doi.org/10.1016/j.rse.2017.01.005>.
- Bratman, G.N., Hamilton, J.P., Daily, G.C., 2012. The impacts of nature experience on human cognitive function and mental health. *Ann. N. Y. Acad. Sci.* 1249, 118–136. <https://doi.org/10.1111/j.1749-6632.2011.06400.x>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* <https://doi.org/10.1023/A:1010933404324>.
- Cai, D., He, X., Han, J., 2007. Semi-supervised discriminant analysis, in: Proceedings of the IEEE International Conference on Computer Vision. 222–228. <https://doi.org/10.1109/ICCV.2007.4408856>.
- Cao, X., Chen, J., Imura, H., Higashi, O., 2009. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sens. Environ.* 113, 2205–2209. <https://doi.org/10.1016/j.rse.2009.06.001>.
- Chen, Z., Xu, B., Devereux, B., 2014. Urban landscape pattern analysis based on 3D landscape models. *Appl. Geogr.* 55, 82–91. <https://doi.org/10.1016/j.apgeog.2014.09.006>.
- Chen, Z., Yu, B., Song, W., Liu, H., Wu, Q., Shi, K., Wu, J., 2017. A new approach for detecting urban centers and their spatial structure with nighttime light remote sensing. *IEEE Trans. Geosci. Remote Sens.* 55, 6305–6319. <https://doi.org/10.1109/TGRS.2017.2725917>.
- Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C., Ghosh, T., 2017. VIIRS night-time lights. *Int. J. Remote Sens.* 38, 5860–5879. <https://doi.org/10.1080/01431161.2017.1342050>.
- Groeneveld, R.A., Meeden, G., 1984. Measuring skewness and kurtosis. *J. R. Stat. Soc. Ser. D* 33, 391–399.
- Hale, J.D., Davies, G., Fairbrass, A.J., Matthews, T.J., Rogers, C.D.F., Sadler, J.P., 2013. Mapping lightscape: spatial patterning of artificial lighting in an urban landscape. *PLoS ONE* 8. <https://doi.org/10.1371/journal.pone.0061460>.
- Hu, S., Wang, L., 2013. Automated urban land-use classification with remote sensing. *Int. J. Remote Sens.* 34, 790–803. <https://doi.org/10.1080/01431161.2012.714510>.
- Hu, T., Yang, J., Li, X., Gong, P., 2016. Mapping urban land use by using Landsat images and open social data. *Remote Sens.* 8 <https://doi.org/10.3390/rs8020151>.
- Huang, B., Zhao, B., Song, Y., 2018a. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens. Environ.* 214, 73–86. <https://doi.org/10.1016/j.rse.2018.04.050>.
- Huang, X., Chen, H., Gong, J., 2018b. Angular difference feature extraction for urban scene classification using ZY-3 multi-angle high-resolution satellite imagery. *ISPRS J. Photogramm. Remote Sens.* 135, 127–141. <https://doi.org/10.1016/j.isprsjprs.2017.11.017>.
- Huang, X., Wang, Y., 2019. Investigating the effects of 3D urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: a case study of Wuhan, Central China. *ISPRS J. Photogramm. Remote Sens.* 152, 119–131. <https://doi.org/10.1016/j.isprsjprs.2019.04.010>.
- Huang, X., Wen, D., Li, J., Qin, R., 2017. Multi-level monitoring of subtle urban changes for the megacities of China using high-resolution multi-view satellite imagery. *Remote Sens. Environ.* 196, 56–75. <https://doi.org/10.1016/j.rse.2017.05.001>.
- Huang, X., Zhang, L., 2010. Comparison of vector stacking, multi-SVMs fuzzy output, and multi-SVMs voting methods for multiscale VHR urban mapping. *IEEE Geosci. Remote Sens. Lett.* 7, 2. <https://doi.org/10.1109/LGRS.2009.203256>.
- Huang, X., Zhang, L., 2011. A multidirectional and multiscale morphological index for automatic building extraction from multispectralgeeye-1 imagery. *Photogramm. Eng. Remote Sensing* 77, 721–732. <https://doi.org/10.14358/PERS.77.7.721>.
- Huang, X., Zhu, T., Zhang, L., Tang, Y., 2014. A novel building change index for automatic building change detection from high-resolution remote sensing imagery. *Remote Sens. Lett.* 5, 713–722. <https://doi.org/10.1080/2150704X.2014.963732>.
- Huang, X., Zhu, Z., Li, Y., Wu, B., Yang, M., 2018c. Tea garden detection from high-resolution imagery using a scene-based framework. *Photogramm. Eng. Remote Sensing* 84, 723–731. <https://doi.org/10.14358/pers.84.11.723>.
- Kane, K., Tuccillo, J., York, A.M., Gentile, L., Ouyang, Y., 2014. A spatio-temporal view of historical growth in Phoenix, Arizona, USA. *Landsc. Urban Plan.* 121, 70–80. <https://doi.org/10.1016/j.landurbplan.2013.08.011>.
- Katz, Y., Levin, N., 2016. Quantifying urban light pollution – a comparison between field measurements and EROS-B imagery. *Remote Sens. Environ.* 177, 65–77. <https://doi.org/10.1016/j.rse.2016.02.017>.
- Klopp, J.M., Petretta, D.L., 2017. The urban sustainable development goal: indicators, complexity and the politics of measuring cities. *Cities* 63, 92–97. <https://doi.org/10.1016/j.cities.2016.12.019>.
- Kuechly, H.U., Kyba, C.C.M., Ruhtz, T., Lindemann, C., Wolter, C., Fischer, J., Hölker, F., 2012. Aerial survey and spatial analysis of sources of light pollution in Berlin, Germany. *Remote Sens. Environ.* 126, 39–50. <https://doi.org/10.1016/j.rse.2012.08.008>.
- Kwon, Y., Kim, S., Jeon, B., 2014. Unraveling the factors determining the redevelopment of Seoul's historic hanoks. *Habitat Int.* 41, 280–289. <https://doi.org/10.1016/j.habitatint.2013.09.003>.
- Levin, N., Duke, Y., 2012. High spatial resolution night-time light images for demographic and socio-economic studies. *Remote Sens. Environ.* 119, 1–10. <https://doi.org/10.1016/j.rse.2011.12.005>.
- Levin, N., Kyba, C.C.M., Zhang, Q., Sánchez de Miguel, A., Román, M.O., Li, X., Portnov, B.A., Molthan, A.L., Jechow, A., Miller, S.D., Wang, Z., Shrestha, R.M., Elvidge, C.D., 2020. Remote sensing of night lights: a review and an outlook for the future. *Remote Sens. Environ.* 237 <https://doi.org/10.1016/j.rse.2019.111443>.
- Li, C., Liu, M., Hu, Y., Shi, T., Qu, X., Walter, M.T., 2018. Effects of urbanization on direct runoff characteristics in urban functional zones. *Sci. Total Environ.* 643, 301–311. <https://doi.org/10.1016/j.scitotenv.2018.06.211>.
- Li, J., Huang, X., Gong, J., 2019a. Deep neural network for remote-sensing image interpretation: status and perspectives. *Natl. Sci. Rev.* 6, 1082–1086. <https://doi.org/10.1093/nsr/nwz058>.
- Li, X., Liu, Z., Chen, X., Sun, J., 2019b. Assessing the ability of LuoJia 1–01 imagery to detect feeble nighttime lights. *Sensors* 19, 1–16. <https://doi.org/10.3390/s19173708>.
- Li, X., Wang, T., Zhang, G., Jiang, B., Jia, P., Zhang, Z., Zhao, Y., 2019c. Planar block adjustment for China's land regions with LuoJia1-01 nighttime light imagery. *Remote Sensing* 11, 2097. <https://doi.org/10.3390/rs11182097>.
- Li, X., Zhou, Y., 2017. Urban mapping using DMSP/OLS stable night-time light: a review. *Int. J. Remote Sens.* 38, 6030–6046. <https://doi.org/10.1080/01431161.2016.1274451>.
- Liu, C., Huang, X., Zhu, Z., Chen, H., Tang, X., Gong, J., 2019. Automatic extraction of built-up area from ZY3 multi-view satellite imagery: analysis of 45 global cities. *Remote Sens. Environ.* 226, 51–73. <https://doi.org/10.1016/j.rse.2019.03.033>.
- Liu, H., Huang, X., Wen, D., Li, J., 2017a. The use of landscape metrics and transfer learning to explore urban villages in China. *Remote Sens.* 9, 1–23. <https://doi.org/10.3390/rs9040365>.
- 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. Geogr. Inf. Sci.* 31, 1675–1696. <https://doi.org/10.1080/13658816.2017.1324976>.
- Lu, C., Yang, X., Wang, Z., Li, Z., 2018. Using multi-level fusion of local features for land-use scene classification with high spatial resolution images in urban coastal zones. *Int. J. Appl. Earth Obs. Geoinf.* 70, 1–12. <https://doi.org/10.1016/j.jag.2018.03.010>.
- Mellander, C., Lobo, J., Stolarick, K., Matheson, Z., 2015. Night-time light data: a good proxy measure for economic activity? *PLoS ONE* 10, 1–18. <https://doi.org/10.1371/journal.pone.0139779>.
- Ou, J., Liu, X., Liu, P., Liu, X., 2019. Evaluation of LuoJia 1–01 nighttime light imagery for impervious surface detection: a comparison with NPP-VIIRS nighttime light data. *Int. J. Appl. Earth Obs. Geoinf.* 81, 1–12. <https://doi.org/10.1016/j.jag.2019.04.017>.
- Peng, J., Xie, P., Liu, Y., Ma, J., 2016a. Urban thermal environment dynamics and associated landscape pattern factors: a case study in the Beijing metropolitan region. *Remote Sens. Environ.* 173, 145–155. <https://doi.org/10.1016/j.rse.2015.11.027>.
- Peng, X., Wang, L., Wang, X., Qiao, Y., 2016b. Bag of visual words and fusion methods for action recognition: comprehensive study and good practice. *Comput. Vis. Image Underst.* 150, 109–125. <https://doi.org/10.1016/j.cviu.2016.03.013>.
- Sallis, J.F., Cerin, E., Conway, T.L., Adams, M.A., Frank, L.D., Pratt, M., Salvo, D., Schipperijn, J., Smith, G., Cain, K.L., Davey, R., Kerr, J., Lai, P.C., Mitás, J., Reis, R., Sarmiento, O.L., Schofield, G., Troelsen, J., Van Dyck, D., De Bourdeaudhuij, I., Owen, N., 2016. Physical activity in relation to urban environments in 14 cities worldwide: a cross-sectional study. *Lancet* 387, 2207–2217. [https://doi.org/10.1016/S0140-6736\(15\)01284-2](https://doi.org/10.1016/S0140-6736(15)01284-2).
- Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. *Environ. Res. Lett.* 4 <https://doi.org/10.1088/1748-9326/4/4/044003>.
- Song, J., Lin, T., Li, X., Prishchepov, A.V., 2018a. Mapping urban functional zones by integrating very high spatial resolution remote sensing imagery and points of interest: a case study of Xiamen, China. *Remote Sensing* 10. <https://doi.org/10.3390/rs10111737>.
- Song, J., Tong, X., Wang, L., Zhao, C., Prishchepov, A.V., 2019. Monitoring finer-scale population density in urban functional zones: a remote sensing data fusion approach.

- Landsch. Urban Plan 190, 103580. <https://doi.org/10.1016/j.landurbplan.2019.05.011>.
- Song, X.P., Hansen, M.C., Stehman, S.V., Potapov, P.V., Tyukavina, A., Vermote, E.F., Townshend, J.R., 2018b. Global land change from 1982 to 2016. *Nature* 560, 639–643. <https://doi.org/10.1038/s41586-018-0411-9>.
- Sun, R., Wang, Y., Chen, L., 2018. A distributed model for quantifying temporal-spatial patterns of anthropogenic heat based on energy consumption. *J. Clean. Prod.* 170, 601–609. <https://doi.org/10.1016/j.jclepro.2017.09.1533>.
- Tang, P., Du, P., Lin, C., Guo, S., Qie, L., 2020. A novel sample selection method for impervious surface area mapping using JLI-3B nighttime light and Sentinel-2 imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 1. <https://doi.org/10.1109/JSTARS.2020.3004654>.
- Taubenböck, H., Klotz, M., Wurm, M., Schmieder, J., Wagner, B., Wooster, M., Esch, T., Dech, S., 2013. Delineation of central business districts in mega city regions using remotely sensed data. *Remote Sens. Environ.* 136, 386–401. <https://doi.org/10.1016/j.rse.2013.05.019>.
- UN, 2018. *World Urbanization Prospects: The 2018 Revision*. United Nations, New York. <https://population.un.org/wup/>.
- Voltersen, M., Berger, C., Hese, S., Schmillius, C., 2014. Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level. *Remote Sens. Environ.* 154, 192–201. <https://doi.org/10.1016/j.rse.2014.08.024>.
- Wen, D., Huang, X., Zhang, A., Ke, X., 2019. Monitoring 3D building change and urban redevelopment patterns in inner city areas of Chinese megacities using multi-view satellite imagery. *Remote Sens.* 11. <https://doi.org/10.3390/rs11070763>.
- Wen, D., Huang, X., Zhang, L., Benediktsson, J.A., 2016. A novel automatic change detection method for urban high-resolution remotely sensed imagery based on Multiindex scene representation. *IEEE Trans. Geosci. Remote Sens.* 54, 609–625. <https://doi.org/10.1109/TGRS.2015.2463075>.
- Wu, S.S., Qiu, X., Usery, E.L., Wang, L., 2009. Using geometrical, textural, and contextual information of land parcels for classification of detailed urban land use. *Ann. Assoc. Am. Geogr.* 99, 76–98. <https://doi.org/10.1080/00045600802459028>.
- Xie, Y., Weng, Q., 2016. Updating urban extents with nighttime light imagery by using an object-based thresholding method. *Remote Sens. Environ.* 187, 1–13. <https://doi.org/10.1016/j.rse.2016.10.002>.
- Xue, X., Lin, Y., Zheng, Q., Wang, K., Zhang, J., Deng, J., Abubakar, G.A., Gan, M., 2020. Mapping the fine-scale spatial pattern of artificial light pollution at night in urban environments from the perspective of bird habitats. *Sci. Total Environ.* 702, 134725. <https://doi.org/10.1016/j.scitotenv.2019.134725>.
- Zhang, H., Jing, X.M., Chen, J.Y., Li, J.J., Schwegler, B., 2016. Characterizing urban fabric properties and their thermal effect using quickbird image and landsat 8 thermal infrared (TIR) data: the case of downtown Shanghai. *China. Remote Sens.* 8, 1–20. <https://doi.org/10.3390/rs8070541>.
- Zhang, Q., Schaaf, C., Seto, K.C., 2013. The Vegetation adjusted NTL Urban Index: A new approach to reduce saturation and increase variation in nighttime luminosity. *Remote Sens. Environ.* 129, 32–41. <https://doi.org/10.1016/j.rse.2012.10.022>.
- Zhang, Q., Seto, K.C., 2011. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sens. Environ.* 115, 2320–2329. <https://doi.org/10.1016/j.rse.2011.04.032>.
- Zhang, X., Du, S., Wang, Q., 2018. Integrating bottom-up classification and top-down feedback for improving urban land-cover and functional-zone mapping. *Remote Sens. Environ.* 212, 231–248. <https://doi.org/10.1016/j.rse.2018.05.006>.
- 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. <https://doi.org/10.1016/j.isprsjprs.2017.09.007>.
- Zhang, X., Du, S., Wang, Y.C., 2015. Semantic classification of heterogeneous urban scenes using intrascene feature similarity and interscene semantic dependency. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8, 2005–2014. <https://doi.org/10.1109/JSTARS.2015.2414178>.
- Zheng, Q., Weng, Q., Huang, L., Wang, K., Deng, J., Jiang, R., Ye, Z., Gan, M., 2018. A new source of multi-spectral high spatial resolution night-time light imagery—JLI-3B. *Remote Sens. Environ.* 215, 300–312. <https://doi.org/10.1016/j.rse.2018.06.016>.
- Zhou, G., Li, C., Li, M., Zhang, J., Liu, Y., 2016. Agglomeration and diffusion of urban functions: an approach based on urban land use conversion. *Habitat Int.* 56, 20–30. <https://doi.org/10.1016/j.habitatint.2016.04.002>.
- Zhou, W., Ming, D., Lv, X., Zhou, K., Bao, H., Hong, Z., 2020. SO-CNN based urban functional zone fine division with VHR remote sensing image. *Remote Sens. Environ.* 236. <https://doi.org/10.1016/j.rse.2019.111458>.
- Zhu, Q., Zhong, Y., Wu, S., Zhang, L., Li, D., 2018. Scene classification based on the sparse homogeneous-heterogeneous topic feature model. *IEEE Trans. Geosci. Remote Sens.* 56, 2689–2703. <https://doi.org/10.1109/TGRS.2017.2781712>.