ISPRS Journal of Photogrammetry and Remote Sensing 198 (2023) 197–209

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Developing an intelligent cloud attention network to support global urban green spaces mapping

Yang Chen [a,b], Qihao Weng [c,*], Luliang Tang [b,*], Lei Wang [d], Hanfa Xing [a], Qinhuo Liu [e]

[a] *Beidou Research Institute, Faculty of Engineering, South China Normal University, Foshan 528000, China*
[b] *State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China*
[c] *Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
[d] *Key Lab for Garden Plant Germplasm Development & Landscape Eco-restoration in Cold Regions of Heilongjiang Province, Northeast Forestry University, Harbin 150040, China*
[e] *State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China*

ARTICLE INFO

Keywords:
Urban green spaces
Urbanization
Sustainable development goals
Cloud removal
Cloud attention intelligent network
Harmonized Landsat-8 and Sentinel-2 data

ABSTRACT

Urban green spaces (UGS) play an important role in understanding of urban ecosystems, climate, environment, and public health concerns. Satellite derived UGS maps provide an efficient and effective tool for urban studies and contribute to targets and indicators of the sustainable development goals, at the global level, set by the United Nations. However, clouds create a challenging issue in optical satellite image processing, leading to significant uncertainty in UGS mapping. In this study, we propose an automatic UGS mapping method by integrating satellite images with crowdsourced geospatial data while aiming to reduce the uncertainty caused by cloud contamination. The proposed method consists of three parts: (1) auxiliary data pre-processing module; (2) cloud attention intelligent network (CAI-net); and (3) non-cloud scenes classification module. The auxiliary data pre-processing module was used to convert crowdsourcing geospatial data into auxiliary maps. The CAI-net was proposed to retrieve detailed UGS classes within clouds from satellite image patches and auxiliary maps, while non-cloud scenes classification module was used to extract UGS from satellite image patches. The proposed method was applied to generate spatial continuous global UGS map products, considering the uncertainty caused by cloud contamination. The results show the proposed method yielded a high-quality global UGS map with average overall accuracy as high as 92.96% when satellite images had cloud coverage ranging from 0% to 50%. The geospatial AI, specifically CAI-net, can provide more accurate UGS mapping regardless of different geographical and climatic conditions of the study areas, which is especially significant for humid tropical and subtropical regions with frequent clouds and rains.

1. Introduction

Urban green spaces (UGS), such as urban trees, grassland, wetland, and water bodies, are important for urban ecosystem services (Gong et al., 2020), human health (Lee et al., 2015), and sustainable city development (Vaidya and Chatterji, 2020). Vegetation in UGS can regulate urban climate by providing shade and facilitating an increase in evapotranspiration (Venkataramanan et al., 2019; Weng, 2012). With rapid population growth, about 68% of the world population will reside in the urban areas by 2050 (UNDESA, 2019). Understanding the spatial patterns of UGS is crucial from the perspectives of both urban ecosystem services and sustainable urban development.

Although the spatial distribution of UGS can be acquired by manual measurements, it is costly and time-consuming (Voltersen et al., 2014). Remote sensing techniques have been widely used for urban land use and cover mapping at city, nation, and global scales (Dong et al., 2017). With the development of moderate-high resolution imaging such as the Landsat, Sentinel-2, and Gaofen-1 series, rapid and accurate UGS mapping has become feasible (Xu and Somers, 2021).

Recently, many UGS mapping methods have been exploited based on optical remote sensing images (Chen et al., 2015; Fu and Qu, 2018; Di et al.,2019; Kranjčić et al. 2019; Liu et al., 2019; Huang et al., 2021). Chen et al. (2015) developed an automatic method integrating pixel and object-based methods with land cover and change knowledge to map global land cover at 30 m resolution. Fu and Qu (2018) used the fully convolutional neural network (FCN) for mapping urban land cover

---

* Corresponding authors.
E-mail addresses: qihao.weng@polyu.edu.hk (Q. Weng), tll@whu.edu.cn (L. Tang).

https://doi.org/10.1016/j.isprsjprs.2023.03.005
Received 8 September 2022; Received in revised form 2 March 2023; Accepted 6 March 2023

Table 1
Information of the experimental data.

Data Name	Sources	Acquisition Date	Spatial Resolution
Sentinel-2A	European Space Agency	2017.1–2017.8	10–60 *m*
HLS data	National Aeronautics and Space Administration	2017.1–2017.8	10–30 *m*
GSWO data	Google Earth Engine	2017.1–2017.8	30 *m*
OSM data	OpenStreetMap platform	2017.8	No

achieving an average overall accuracy of 85%. The object-oriented classification technique was used to extract UGS from RapidEye images (Di et al., 2019). Kranjčić et al. (2019) used SVM method for UGS extraction from Sentinel-2A images. Liu et al. (2019) developed the DeepLabv3plus model and used it for UGS mapping, yielding an overall accuracy of 89.46% from cloud-free GF-2 images. Huang et al. (2021) used random forests classifier based on Google Earth Engine (GEE) platform for mapping UGS from cloud-free Landsat images, and their result indicated that the random forests classifier yielded an average overall accuracy of 89.26%.

Although the abovementioned UGS mapping methods can get relatively high accurate extraction results, these methods used only cloud-free images to generate UGS or related products. However, optical remote sensing images are inevitably contaminated with clouds, especially in humid tropical and subtropical regions, leading to discontinuity in spatial patterns of satellite-derived UGS products (Chen et al., 2022; Chen et al., 2020). It remains a great challenge to derive spatially continuous, quality UGS maps from cloud contaminated satellite images.

To minimize the impact of clouds on optical remote sensing images for urban land use/cover mapping, many researchers have employed multi-temporal images for mapping (Salberg, 2010; Zhao et al.,2016; Sharma et al., 2018; Xie et al., 2019; Jie et al., 2021). Salberg (2010) developed a two-stage classifier for mapping land cover by fusing multi-temporal Landsat images. The random forest method was used to generate land cover maps from multi-temporal Landsat 8 images by Zhao et al. (2016). The patch-based recurrent neural network (PB-RNN) method was used to generate spatially continuous land cover maps from multi-temporal Landsat 8 images (Sharma et al., 2018). Xie et al. (2019) exploited a multi-temporal Landsat classification algorithm to solve the problem of land cover classification in cloud covered areas. Jie et al. (2021) developed a novel two-step multi-temporal image classification method to improve classification accuracy by using multi-temporal images in missing pixel areas. These classification methods utilize multiple cloud free patch images to improve the accuracy of UGS cover classification in cloud contaminated images. However, their performance depends on the availability of cloud-free images. In addition, multi-temporal continuous cloud-free images are difficult to acquire, especially in the tropical and subtropical regions (Chen et al., 2018; Qu et al., 2021).

To solve the issue of the above multi-temporal classification methods, some researchers have attempted to take the advantage of multi-source images fusion, such as Sentinel-2 and Sentinel-1 images fusion, for improving UGS classification accuracy (Sukawattanavijit et al., 2017; Chen et al., 2017; Shrestha et al., 2019; Nuthammachot and Stratoulias, 2019; Hong et al., 2021). Sukawattanavijit et al. (2017) proposed a Ga-SVM algorithm by integrating SAR and multispectral images for improving land-cover classification. Shrestha et al. (2019) exploited a classification method with the fusion of Sentinel-2 and Sentinel-1 images to solve the issue of land cover map in cloud covered areas. Landsat based urban land cover mapping have been extensively conducted by integrating multi-source remote sensing features, including temporal, spectral, and topographic features, to reduce misclassifications of different land cover types (Chen et al., 2017; Nuthammachot and Stratoulias, 2019; Hong et al., 2021). These methods used SAR images for circumventing cloud cover problem (Hang et al., 2020).

Multi-source remote sensing image fusion can provide some auxiliary information for improving urban land mapping, few studies have investigated the efficacy for crowdsourcing geospatial data to improve the accuracy of UGS mapping in cloud-covered regions (Zhang et al., 2017; Johnson et al., 2017). The crowdsourcing geographic data, such as point of interest (POI) data, can provide geo-referenced prior information (Du et al., 2015). The prior information has been widely used to understand UGS functions from a human cognition perspective (Chen et al., 2018). The OpenStreetMap (OSM) platform can provide freely available crowdsourcing geospatial data (Jiang et al., 2015). The integration of remote sensing imagery with crowdsourcing geospatial data can effectively improve the accuracy of urban land cover/use mapping because the crowdsourcing geospatial data provided rich prior information (Su et al., 2020; Chen et al., 2021).

In terms of UGS mapping, Chen et al. (2021) developed a geospatial neural network by integrating remotely sensed imagery with crowdsourced spatial data to map UGS in the Wuhan city. Although this method can achieve high mapping accuracy for a single city, but the geospatial neural network method contained two limitations: (1) It cannot generate spatially continuous UGS maps in cloud covered regions; and (2) while it can generate UGS map for a single city, it cannot be applied to other cities without label sample verification.

In this study, the UGS were referred to urban vegetation, wetland, water bodies and other kinds of natural environment. This study exploits the integration of remote sensing images and crowdsourcing geospatial big data for improving UGS mapping in cloud-covered areas at the global level. The objectives of this study are: (1) To generate spatially and temporally continuous UGS maps in cloud covered areas; (2) To generate a high-quality global UGS data product. We propose an automatic UGS mapping method by integrating satellite images with crowdsourced geospatial data. The proposed method consists of three parts: (1) auxiliary data pre-processing module; (2) cloud attention intelligent network (CAI-net); and (3) non-cloud scenes classification module. Cloud cover is an inevitable issue in optical satellite data, causing mapping uncertainty and spatial discontinuity in UGS maps. Inspired by the attention mechanism in deep learning (Vaswani et al., 2017), we have designed a cloud attention intelligent network (CAI-net) integrated Sentinel-2 images with crowdsourced geospatial data to eliminate cloud effects in spatial–temporal continuous UGS mapping. Due to the geospatial neural network achieved excellent UGS mapping results at non-cloud scenes (Chen et al., 2021), we used the state-of-the-art geospatial neural network for extracting UGS in this study.

2. Data

We selected freely available the satellite and auxiliary data as the data sources for mapping global UGS. The satellite and auxiliary data included Sentinel-2A images, Harmonized Landsat-8 and Sentinel-2 data (HLS), OSM data, Global Surface Water Occurrence (GSWO). The satellite and ancillary data description are listed in Table 1.

2.1. Satellite images

The Sentinel-2A images acquired by MultiSpectral Instrument (MSI), have been widely used for monitoring global vegetation, water resources, urban land changes, and ecosystem service mapping (Punalekar et al., 2018). The Sentinel-2A were freely downloaded from the European Space Agency (), which has 13 spectral bands (spectrum range from 430 to 2320 nm) high time resolution (revisit period of 5 days), and high spatial resolution (spatial resolution range from 10 m to 60 m). The Sentinel-2A imagery have been found suitable for urban mapping since it can offer global surface coverage with high spatial resolution (Preidl et al., 2020). The Sentinel-2A (L2A level) images were adopted for

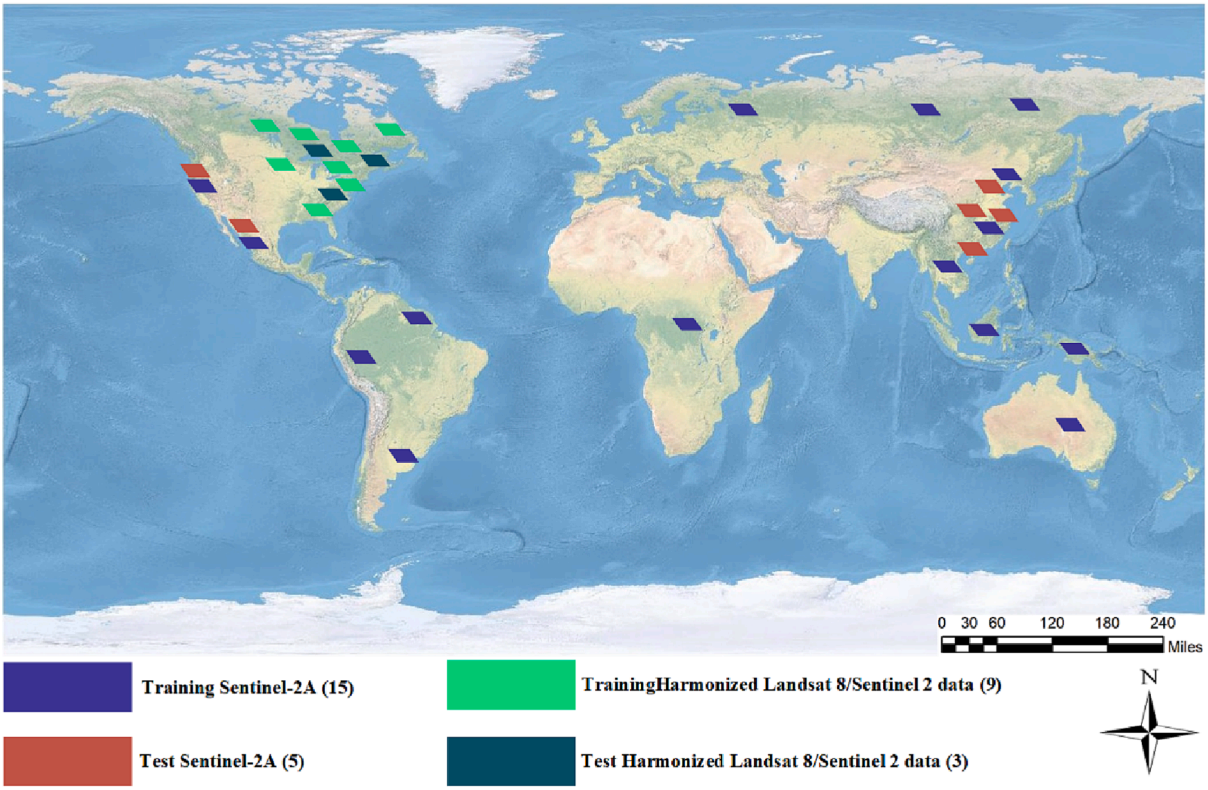

Fig. 1. Global distribution of satellite image acquisition for the study, including Sentinel-2A images and Harmonized Landsat and Sentinel-2 data.

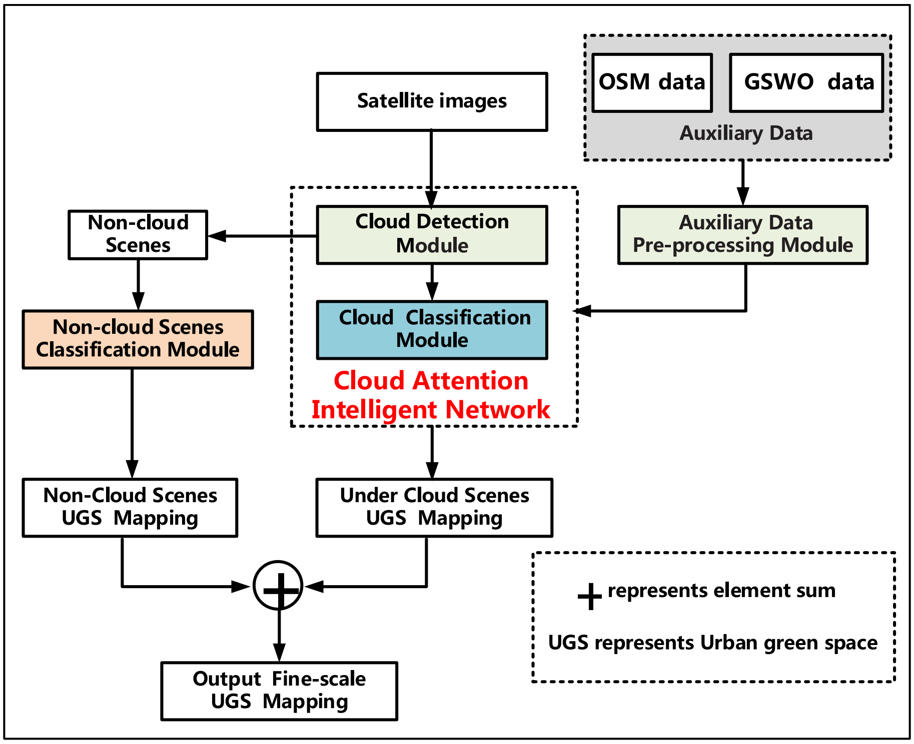

Fig. 2. The overall workflow of the proposed method with cloud attention intelligent network (CAI-net).

mapping spatial continuous UGS. The atmospheric correction and radiometric calibration were implemented by ESA Sentinel Application Platform (SNAP) and ENVI 5.6 Platform.

Combining data from Landsat OLI and Sentinel-2 MSI data can provide a new resource for Earth observation because the harmonized Landsat-8 and Sentinel-2 (HLS) data reduces global observation of the land to every 2–3 days (Punalekar et al., 2018). It was produced by National Aeronautics and Space Administration (NASA) project, which

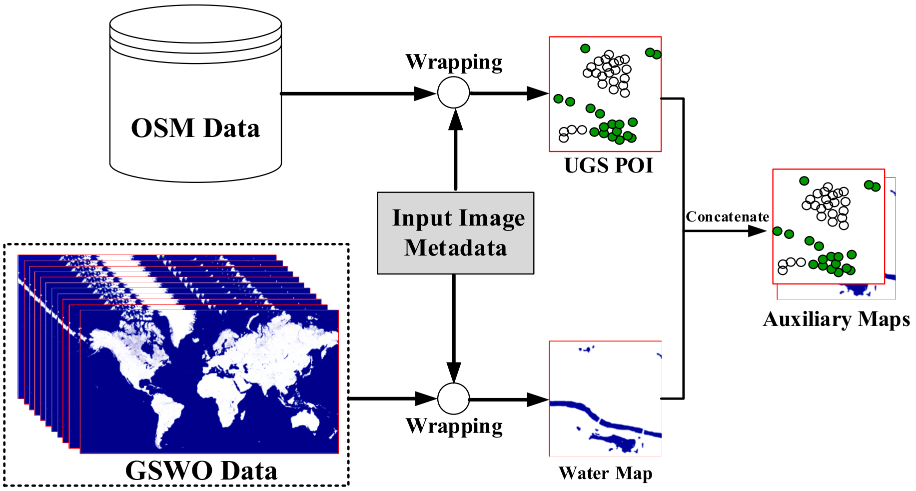

Fig. 3. The processing procedure for the auxiliary data.

Table 2
UGS classes derived from the initial POI types.

UGS classes	Initial POI types
UGS	urban forest, rivers, urban wetland, green belts, park, residential green space, botanical garden natural reserve
Non-UGS	residence districts, stations, supermarket, airports, buildings, industrial districts, sports center, village green, etc.

intended to provide dense time series of reflectance data (Bolton et al., 2020). The HLS datasets have been widely used for studies of land use mapping, land surface phenology dynamics, and urban mapping tasks (Chaves et al., 2020; Moon et al., 2021; Roy et al., 2019).

The HLS Version 2 datasets have obtained seamless surface reflectance data from Landsat-8 and Sentinel-2 by the harmonization process, which consisted of a set of operations: atmospheric correction, cloud process, spectral band pass adjustments, and spatial co-registration (Claverie et al., 2018). The HLS datasets contained of three products: (1) S10: full resolution MSI (10 m, 20 m and 60 m); (2) S30: MSI harmonized surface reflectance resampled to 30 m; (3) L30: OLI harmonized surface reflectance and TOA brightness temperature resampled to 30 m (Claverie et al., 2018). Considering the differences in vegetation phenology in different climatic zones, we selected Sentinel-2A and HLS images from those zones between 1st August 2017 and 30th August 2017 as training datasets. Fig. 1 shows the global distribution of acquisition location of satellite images.

2.2. Auxiliary data

The OSM platform can provide rich volunteered geographical information such as POI category, POI place, POI time information, and so on (Johnson and Iizuka, 2016). Therefore, the OSM platform shows an important potential for providing prior information to support satellite image analysis tasks. The OSM POI data can be freely obtained from the OpenStreetMap platform (https://www.openstreetmap.org.). To reduce time interval between OSM POI data and Sentinel-2A images, the POI acquisition date was set to 15th August 2017, when shows the distribution of OSM data with 159,993 POI altogether covering the globe. The UGS POI was extracted from OSM POI data. To retrieve water bodies of UGS classes in cloud regions, Global Surface Water Occurrence (GSWO) data were used as auxiliary data. The GSWO were derived from Landsat satellite images (Pekel et al., 2016). GSWO dataset provides long-term water-surface maps for the entire globe and can be freely acquired in GEE platform (https://global-surface-water.appspot.com/map).

3. Methods

We designed a detailed workflow to generate spatially continuous UGS products (see Fig. 2). The proposed method is comprised of three parts: (1) auxiliary data pre-processing module; (2) CAI-net; and (3) non-cloud scenes classification module. The auxiliary data pre-processing module was used to convert the crowdsourcing geospatial data into auxiliary maps. The CAI-net was proposed to retrieve detailed UGS classes in cloud regions from satellite image patches and auxiliary maps. Non-cloud scenes classification module was used to extract UGS from satellite image patches in cloud-free regions.

3.1. Auxiliary data Pre-processing module

We have selected the POI and GSWO data as auxiliary data due to POI and GSWO data can provide the prior UGS features in the globe (Chen et al., 2018). In this study, an auxiliary data pre-processing module was designed to encode two types of crowdsourcing geospatial data, i.e., POI and GSWO data, into a group of auxiliary maps. Fig. 3 shows the flowchart of the auxiliary data pre-processing.

The POI data often include vegetation, parks, residential points, urban wetland, hotels, railway stations, and so on. The crowdsourcing geospatial datasets can describe characteristics of different UGS types (Heikinheimo et al., 2020). To identify UGS classes within cloud covered areas, the crowdsourcing geospatial data were preprocessed as auxiliary information.

The UGS POI density map was the auxiliary information to reflect the UGS existence probability in cloud covered regions. For POI data pre-processing, the POI classes were categorized into the UGS classes. The rules of grouping are introduced in Table 2. UGS POI density map A_{POI} was generated with the kernel density estimation by ArcGIS 10.4. The kernel bandwidth was of the kernel density estimation was set to 10 m. For GSWO data pre-processing, the first longitude-latitude range can be obtained from the input image metadata. The auxiliary water map A_{water} was generated by using the method of pixel cropping. The strategy of aggregation was adopted to integrate an auxiliary map of the POI data and GSWO data. After obtaining pre-processing of auxiliary data, the input image with size $X \times Y$ was selected as the current UGS mapping object. The final auxiliary map A can be defined using the following equation:

$$A = concat(A_{POI}, A_{Water}) \tag{1}$$

where the dimension of the final auxiliary map A is $(X, Y, 2)$.

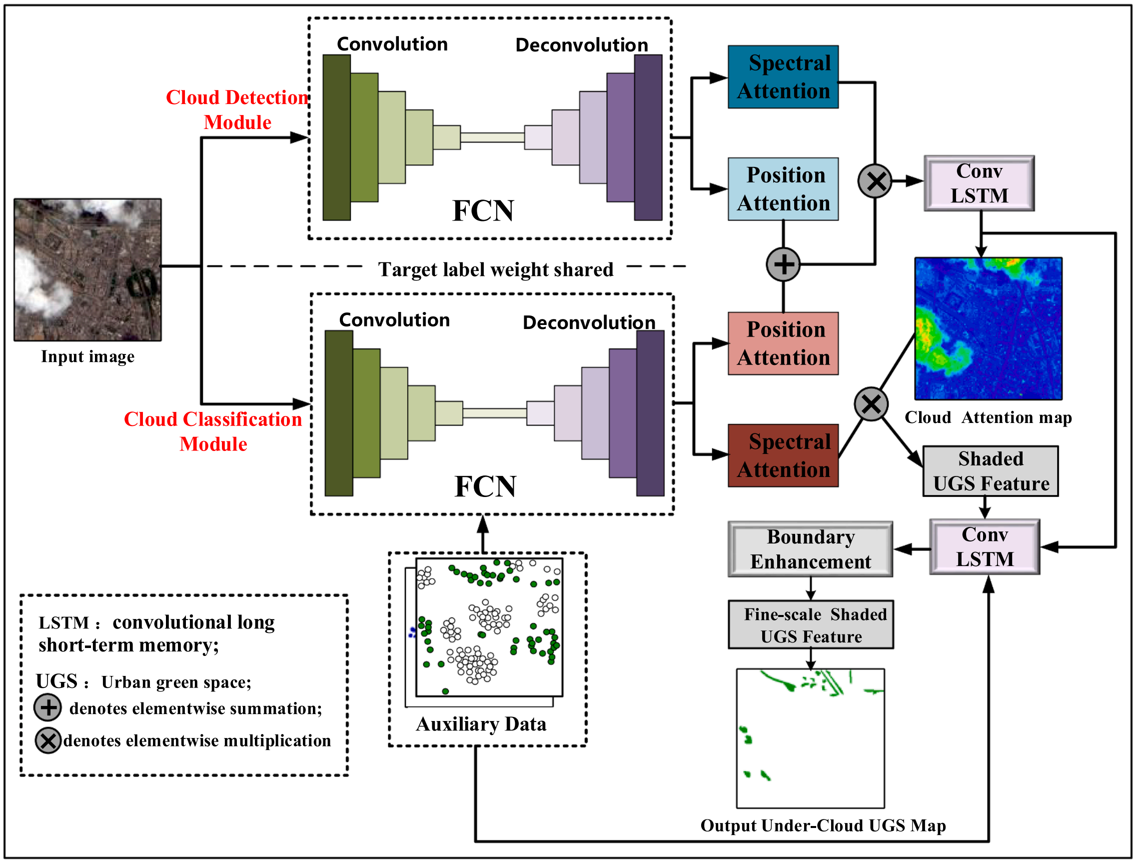

Fig. 4. The architecture of the proposed cloud attention intelligent network.

3.2. The cloud attention intelligent network

The cloud attention intelligent network (CAI-net) retrieve UGS classes in cloud covered areas. The CAI-net is comprised of two modules: (1) cloud detection module (CDM), and (2) cloud classification module (CCM). The CDM was designed to focus cloud position on input cloud contaminated images. The CCM was designed to recognize UGS classes from remote sensing images and crowdsourcing geospatial data in cloud covered areas. The two sequential convolutional long short-term memory (LSTM) were used to refine the cloud and the clouded UGS feature attention in CDM and CCM. Fig. 4 shows the architecture of the proposed CAI net.

3.2.1. Cloud detection module (CDM)

Cloud detection is an essential step for identifying UGS classes in cloud contaminated images. Since fully convolutional networks (FCN) show state-of-the-art performance in the semantic segmentation field (Long et al., 2015), the CDM used an FCN to extract the deep local–global cloud feature, which was used to identify cloud and non-cloud regions in Sentinel-2A images. We proposed a novel CDM architecture with FCN and the convolutional block attention module (Woo et al., 2018), an effective attention network that can integrate the position and the spectral attention. The convolutional block attention module (CBAM) was used in the CDM architecture for its effectiveness in refining cloud contextual semantic features (Zhang et al., 2021). Two unique operations of CDM were applied: (i) The spectral attention was used to learn 'what' was important in cloud contaminated images by exploiting the inter-spectral relationship of cloud information; and (ii) the position attention was used to highlight the positions of the cloud regions, while the spectral attention selected the important bands.

The encoder of CDM executed feature extraction using FCN to produce a set of cloud feature maps. The cloud feature maps were unsampled by a set of deconvolution kernels in the decoder of CDM. Prior to the two attention modules sequence, we used a convolution kernel of size 1 × 1 to adjust the spectral dimension of the features from spectral attention block and position attention block. In the CDM architecture, the sequential algorithm was used to integrate spectral and position attention (Liu and Deng., 2015), which can aggregate local–global semantics cloud information. The CDM architecture can adaptively highlight on more important information and suppresses less important information in cloud contaminated images.

3.2.2. Cloud classification module (CCM)

Prior information was obtained from auxiliary data (such as crowdsourced geospatial data), which can describe the UGS class features (Heikinheimo et al., 2020) and plays an important role in retrieving UGS classes in cloud contaminated areas (Johnson et al., 2017). Therefore, in this paper, we used prior information as auxiliary information to retrieve UGS classes in cloud covered areas. To obtain UGS prior information, the cloud classification module was designed. The feature images from CCM are up-sampled to the size of the input image by using deconvolution operations. Then, the spectral dimension of the up-sampled feature images was adjusted from spectral attention block and position attention block to the same number of spectrums of the input image by using 1 × 1 convolution kernel. The final concatenated feature maps F from the CCM can be defined by the following equation:

$$F = concat(F_{img,1}, F_{img,2}, F_{img,3}, F_{img,4}, F_{aux,1}, F_{aux,2}, F_{aux,3}, F_{aux,4}) \tag{2}$$

where the $F_{img,1}, F_{img,2}, F_{img,3}, F_{img,4}$ are the features from the four spectral bands (i.e., Sentinel-2A Blue band, Green band, Red band, and NIR band).$F_{aux,1}, F_{aux,2}, F_{aux,3}, F_{aux,4}$ features from the auxiliary image.

Compared to the cloud detection module, the cloud classification module has unique operations, including: (i) extracting the UGS features

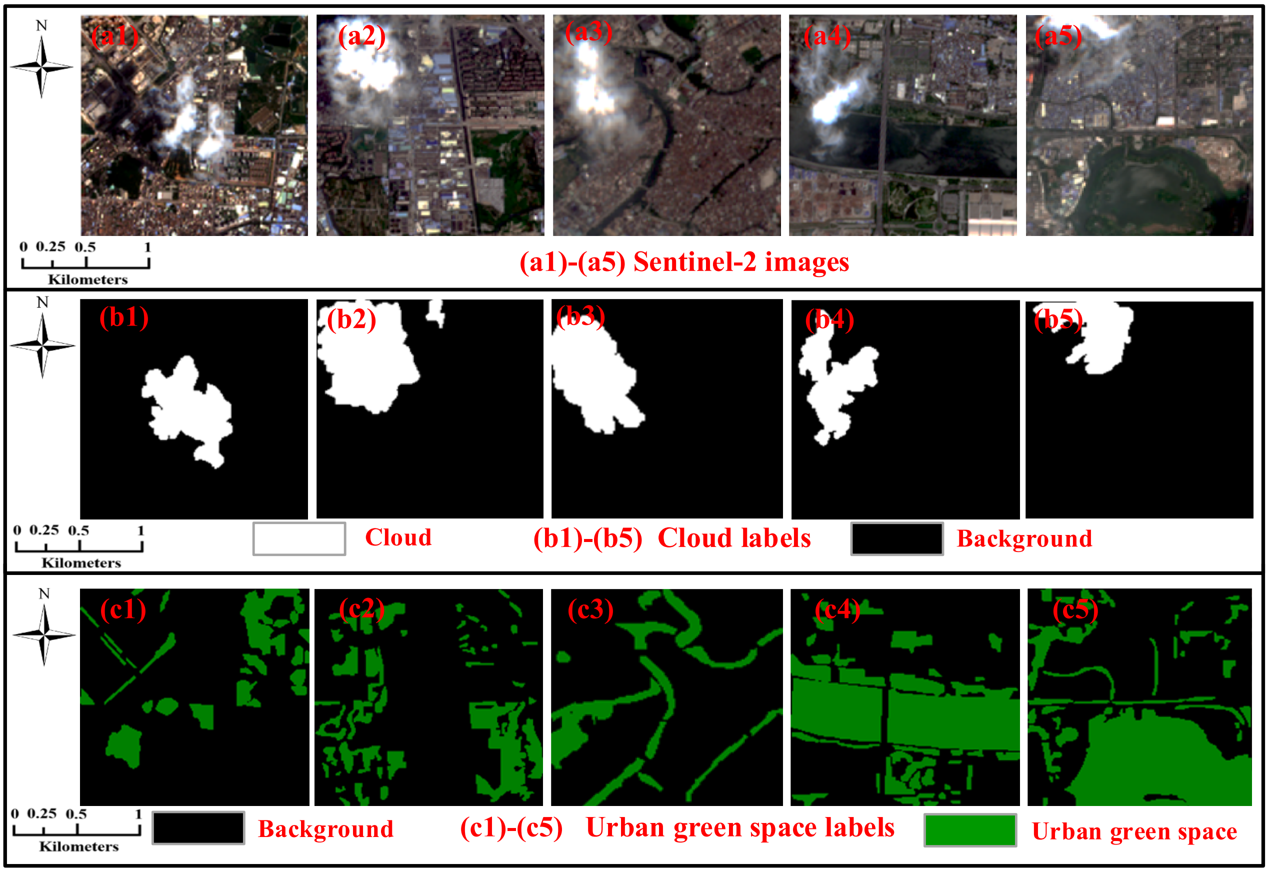

Fig. 5. Some sample patches of urban areas.

Table 3
Major input parameters of the four networks.

Input parameters	Training sample	Mini-batch size	Momentum	Learning rate	Epochs
DeepLab v3plus	70,000	128	0.9	0.01	1000
FCN	70,000	128	0.9	0.01	1000
Geospatial neural network	70,000	128	0.9	0.01	1000
Proposed method	70,000	128	0.9	0.01	1000

for cloud contaminated areas based on the auxiliary feature maps; and (ii) updating the cloud attention of the CDM by using an element-wise sum algorithm on the maps of CDM and CCM. At each stage, the two convolutional LSTM networks (Shi et al., 2015) were used to refine the cloud attention map to implement the CAI-net. The boundary enhancement module with three dilated convolution kernels (the convolution kernel sizes of 5 × 5, 3 × 3, and 1 × 1) was further applied to improve the boundaries of the UGS map (Chen et al., 2021b). The integrated cloud attention map was fed into the CDM and the CCM at the next stage to improve the accuracy of cloud detection and thus the classification of UGS classes in cloud contaminated areas.

3.3. Non-cloud scenes classification module

The Geospatial neural network has been successfully applied to achieve the high accuracy of UGS mapping in cloud-free regions (Chen et al., 2021). Therefore, in this research, the Geospatial neural network was used to extract UGS from Sentinel-2A images in cloud-free regions. The final spatial continuous UGS map A can be defined as follows:

$$A = A_{Cloud} + A_{Non-cloud} \tag{3}$$

where A_{Cloud} is UGS map from cloud attention intelligent network in cloud contaminated regions. $A_{Non-cloud}$ is UGS map from non-cloud scenes classification module in cloud-free regions.

In the proposed method, we used a 1 × 1 convolution kernel to generate the score maps of two classes: background M_1 and UGS M_2.We use softmax function (Peng et al., 2017) to normalize the range of score maps, and the softmax classifier to identify UGS from score maps of different classes. The probability map P_s of each class $s = \{1, 2\}$ can be defined as follows:

$$P_s = \frac{\exp(M_s)}{\sum_{n=1}^{2}\exp(M_n)} \tag{4}$$

4. Results

To evaluate the effectiveness of the proposed network, large experiments were conducted to generate UGS maps with Sentinel-2A images in different cities. The proposed network is compared with the state-of-the-art deep learning methods (DeepLab v3plus, FCN, and Geospatial neural network). The proposed network training and experimental results are described in this section.

4.1. Network training and evaluation metrics

Considering vegetation phenology characteristics, we selected Sentinel-2A and HLS images of the summer season as training datasets. The input images were cropped into patch size of 512 × 512 pixels by using the sliding window strategy with stride size of 1. According to predefined land use classification map, the experimental datasets contained the different types of urban scenes, such as urban forest, urban

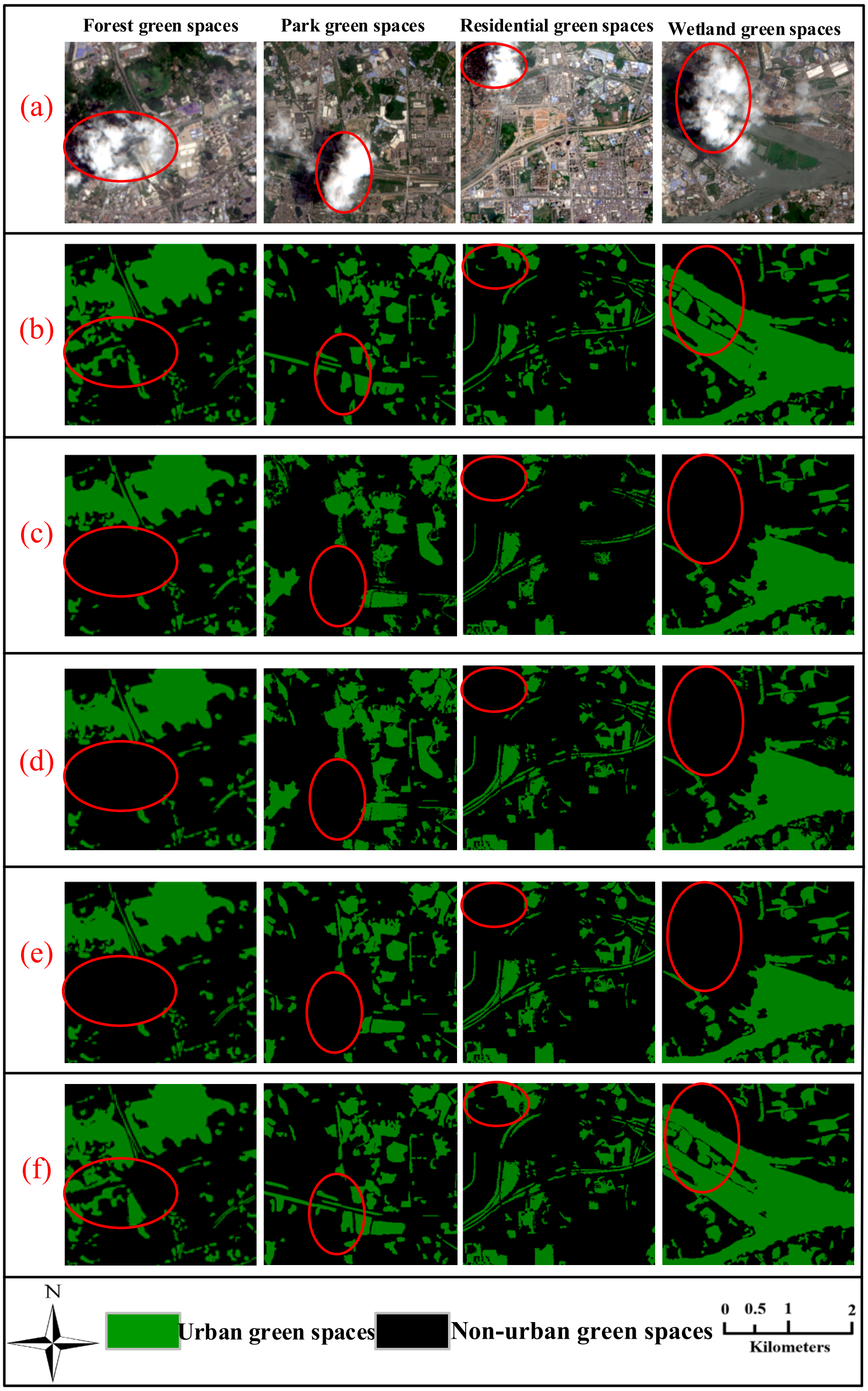

Fig. 6. The results of extracting urban green spaces by different methods for different urban land-cover scenes. (a) Cloud of Sentinel-2 data; (b) The ground truth reference; (c) DeepLab v3plus; (d) FCN; (e) Geospatial neural network; (f) The proposed network. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

park, and commercial district. The UGS and cloud labels were obtained by manual drawing in ENVI platform. Accuracy of manual drawing was controlled within two pixels. Our datasets contained 70,000 training, 15,000 validation, and 15,000 testing image patches.

Fig. 5 show some of the training image patches, cloud labels, and UGS labels. The proposed CAI net was implemented in the PyTorch and Google Earth Engine (GEE) platform, using the stochastic gradient descent to optimize parameters with weight decay of 0.0001,

Table 4
Quantitative evaluation results of the four mapping methods.

Method	OA (%)	Kappa (%)	Recall (%)	Precision (%)
DeepLab v3plus	70.51	66.12	67.31	68.01
FCN	76.34	72.04	72.91	73.39
Geospatial neural network	81.59	74.91	76.97	77.01
Proposed method	95.13	91.09	92.01	93.42

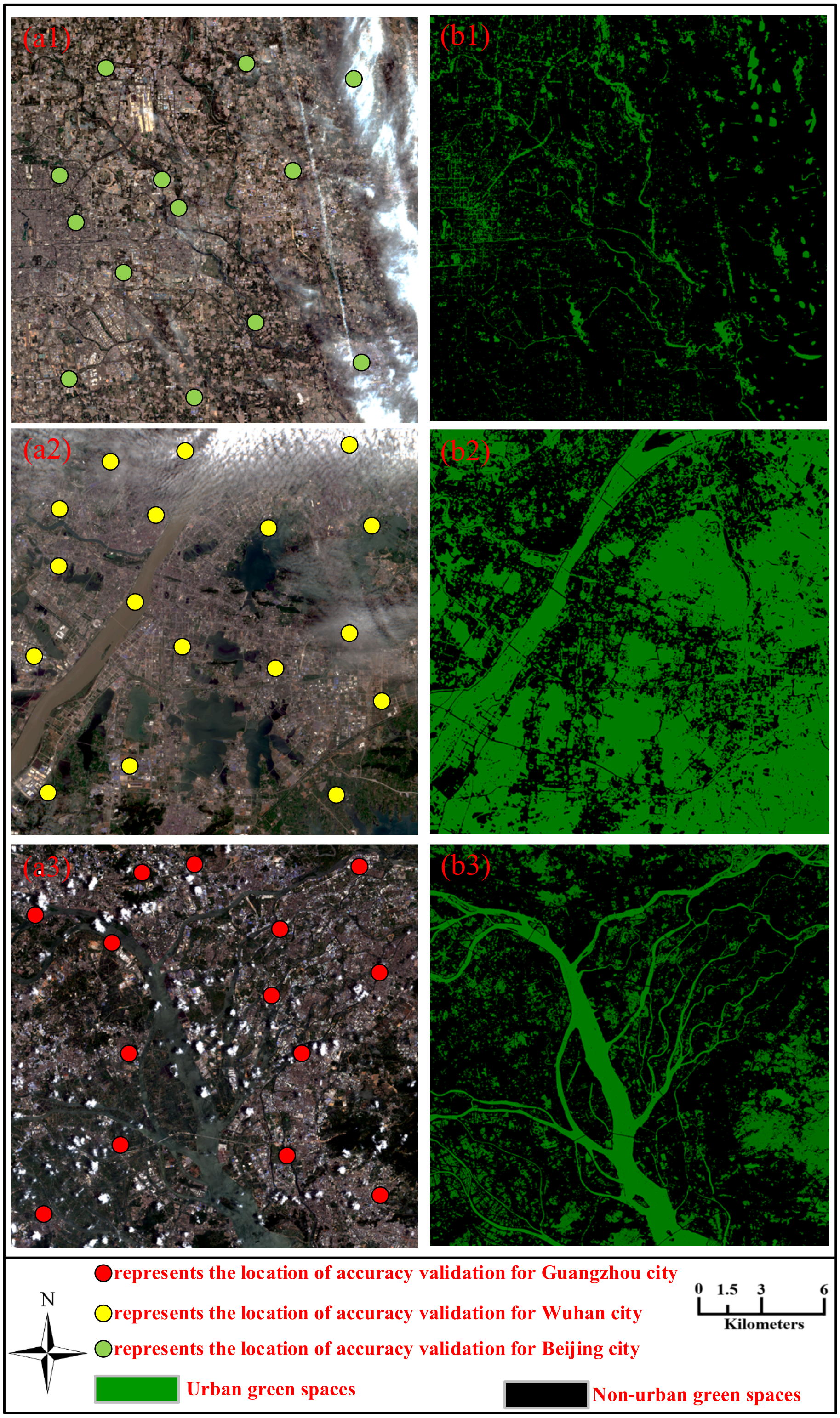

Fig. 7. Urban green spaces mapping for different geographical patterns. (a1)-(a3) Cloud of Sentinel-2A data in Beijing (acquired on August 5, 2016), Wuhan (acquired on August 15, 2016), and Guangzhou (acquired on August 23, 2016), China, respectively; (b1)-(b3) The results of extracted urban green spaces by using the proposed network in Beijing, Wuhan, and Guangzhou, China, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Accuracy assessment of UGS maps using the proposed method for different urban patterns.

City	OA (%)	Kappa (%)	Recall (%)	Precision (%)
Beijing	93.81	90.02	91.18	92.28
Wuhan	95.24	92.38	93.81	94.72
Guangzhou	94.09	90.89	92.31	93.08

momentum of 0.9, and a mini-batch size of 128 for generating spatially continuous UGS maps. The learning rate was initially set to 0.01 and strategy decay by using an exponential decay with decay rate of 0.9 and an epoch of 50 for a total of 1000 epochs in the training. The proposed network and the compared methods were pre-trained by using the same datasets. For FCN-based architectures, the weight parameters of the trained FCN-GoogLeNet (Long et al., 2015) were used as the initially weight of the FCN encoder. All network architectures were implemented

Table 6
Sentinel-2 images over with cloud-cover (0% to 50%) for evaluating CAI net performance.

Cloud Cover	Product ID	Acquisition Date
0%	S2A_MSIL1C_20190817T023551_N0208_R089_T51RUQ_20190817T042607	2019/08/17
5%	S2A_MSIL1C_20190714T025551_N0208_R032_T50SNJ_20190714T055344	2019/07/14
10%	S2A_MSIL1C_20190522T030551_N0209_R075_T50TMK_20190522T060100	2019/05/14
15%	S2A_MSIL1C_20190614T025551_N0207_R032_T49QGF_20190614T055620	2019/06/14
20%	S2A_MSIL1C_20190611T024551_N0207_R132_T50RMS_20190611T052620	2019/06/11
25%	S2A_MSIL1C_20190614T025551_N0207_R032_T49QHF_20190614T055620	2019/06/14
30%	S2A_MSIL1C_20190823T025551_N0208_R032_T50QKL_20190823T061030	2019/08/23
35%	S2A_MSIL1C_20190823T025551_N0208_R032_T50RKV_20190923T061030	2019/09/23
40%	S2A_MSIL1C_20190806T030551_N0208_R075_T49RGQ_20190806T060210	2019/08/15
45%	S2A_MSIL1C_20190803T025551_N0208_R032_T50RLS_20190803T063256	2019/08/03
50%	S2A_MSIL1C_20190724T025551_N0208_R032_T49QHF_20190724T060504	2019/07/24

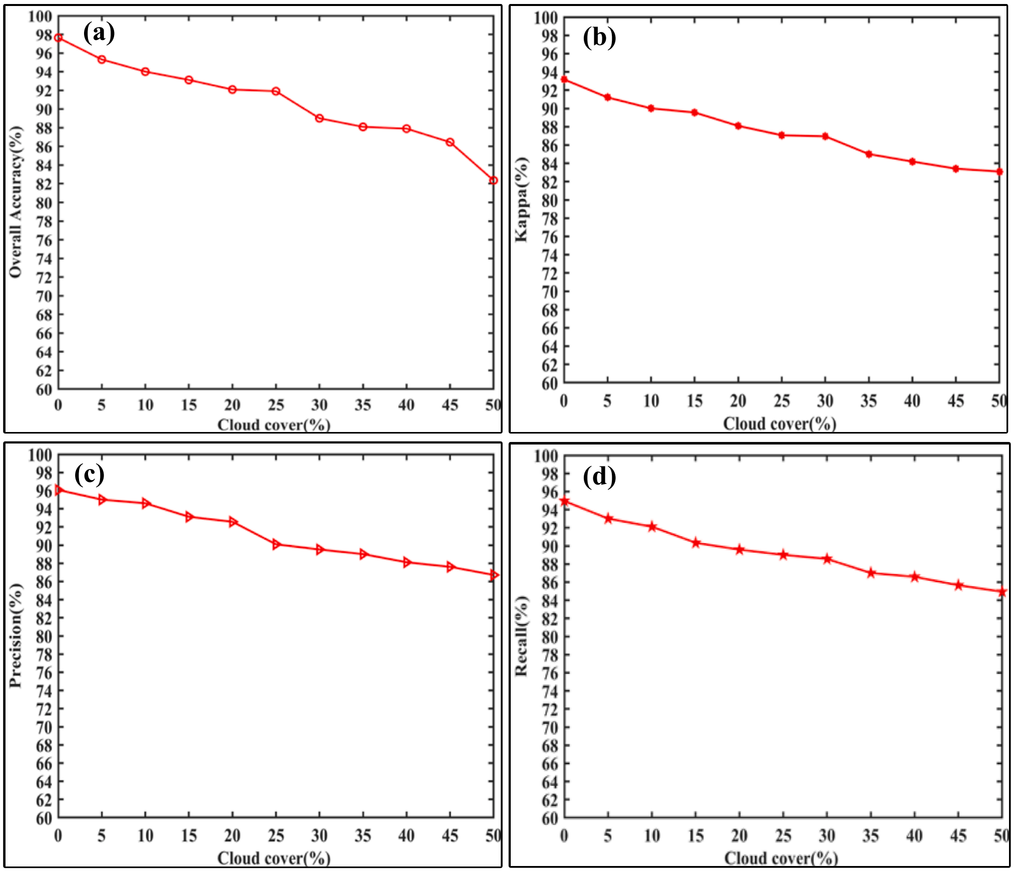

Fig. 8. Effects of the different cloud cover of Sentinel-2 data on urban green spaces mapping accuracies based on 11 Sentinel-2 images. (a) Overall accuracy; (b) Kappa; (c) Precision; (d) Recall. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

on NVIDIA® Quadro M2200 GPU.

The common metrics for assessment for UGS maps include overall accuracy (OA), kappa coefficient (Kappa), recall, and precision (Foody, 2020). They can be defined by following equations, respectively:

$$OA = \frac{(TN + TP)}{N} \tag{5}$$

$$Kappa = \frac{N \times (TN + TP) - [(TP + FP) \times (TP + FN) + (FN + TN) \times (FN + TN)]}{N^2 - [(TP + FP) \times (TP + FN) + (FN + TN) \times (FN + TN)]} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

where N represents the total number of pixels for an entire input image.TP represents the number of UGS pixels accurately extracted.FP represents the number of non-UGS pixels extracted as UGS pixels.TN represents the number of non-UGS pixels accurately extracted. FN represents the number of UGS -pixels extracted as non-UGS pixels.

4.2. The performance of the proposed network

To better assess the performance of the proposed network, we compared it with DeepLab v3plus (Liu et al., 2019), FCN (Long et al., 2015), and Geospatial neural network (Chen et al., 2021b). The state-of-the-art deep learning methods (DeepLab v3plus, FCN, and Geospatial neural network) without cloud detection have been successfully applied to generate UGS map in the urban areas. All the methods used the same input parameters. Input parameters of four networks are listed in Table 3. We initialized DeepLab v3plus, FCN, and Geospatial neural

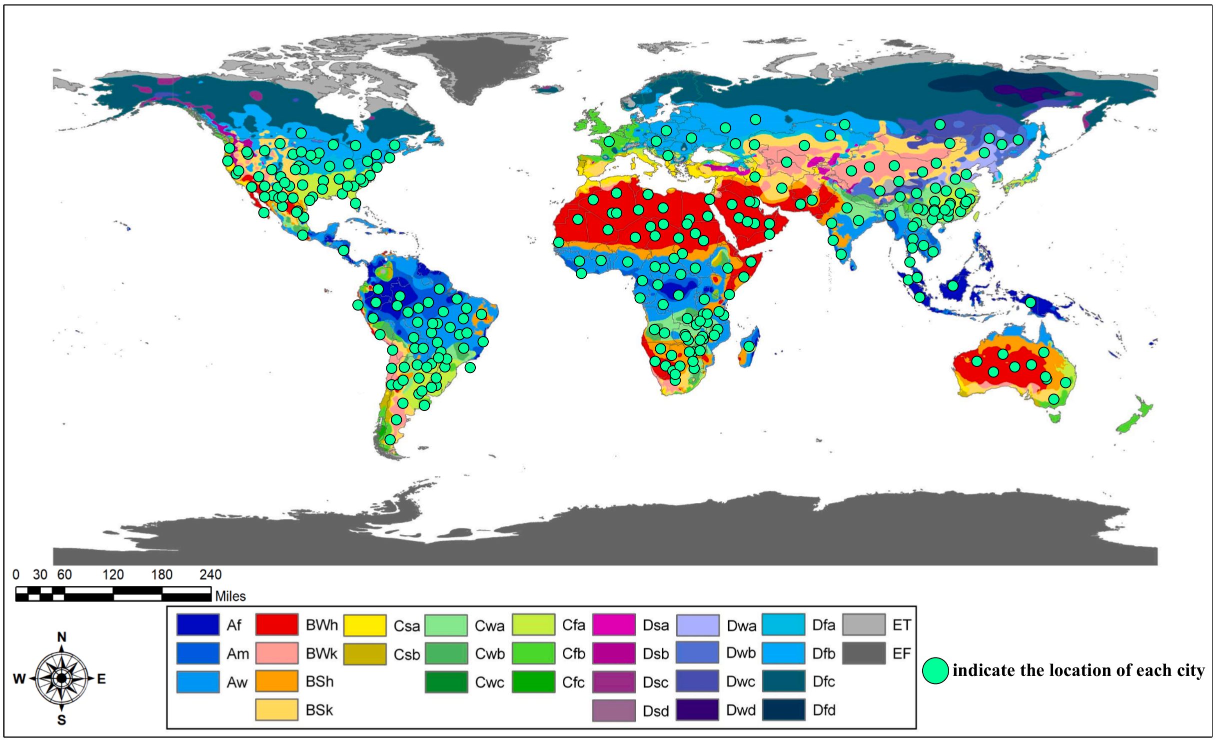

Fig. 9. Distribution of the study cities in different climatic zones (Source of the map: World map of Köppen-Geiger climate zones (Peel et al., 2007).

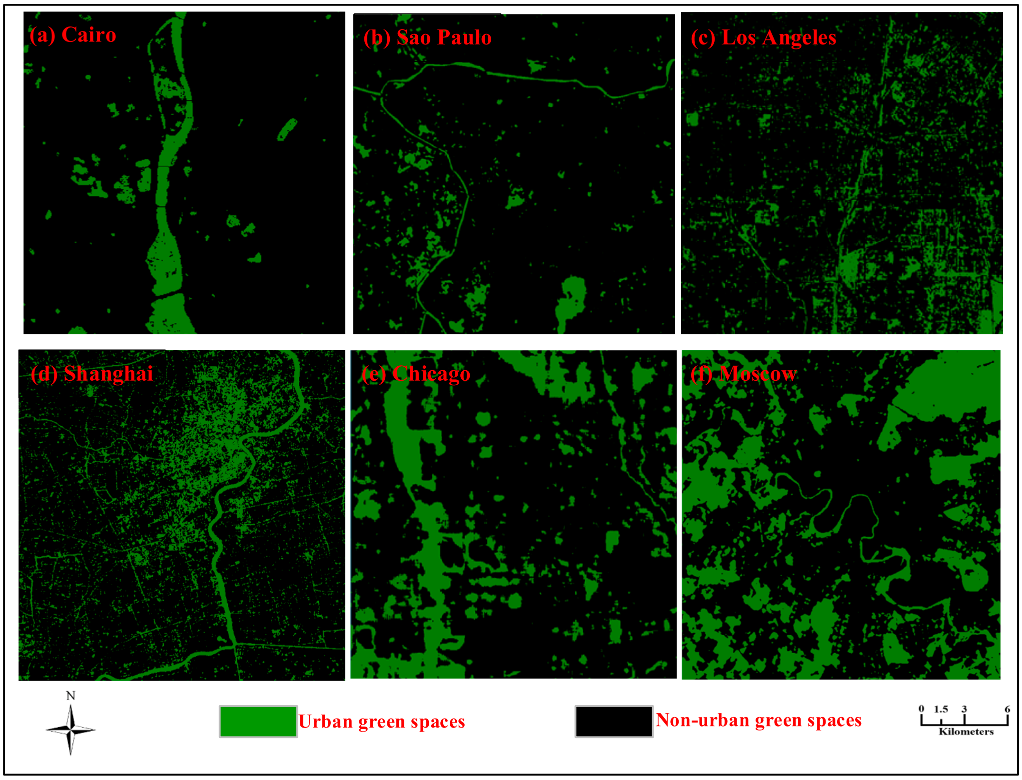

Fig. 10. Examples of the UGS maps in global cities derived from HLS data and Sentinel-2A data using CAI-net.

network, the proposed network, respectively, using the weight parameters from the trained FCN-GoogLeNet. In addition, we have selected the four types of urban land-cover scenes as experimental areas. Four types of urban land-cover scenes contained forest, park green spaces, residential green spaces, and wetlands.

Fig. 6 shows the results of extracting urban green spaces by different methods for different urban land-cover scenes. Fig. 6 (f) shows the proposed network can generate spatially continuous UGS maps from cloudy Sentinel-2 images. DeepLab v3plus, FCN, and Geospatial neural network obtained good UGS extraction results in non-cloud areas;

Table 7
Accuracy assessment of UGS maps of selected global cities in different climatic zones.

City	OA (%)	Kappa (%)	Recall (%)	Precision (%)
Cairo	95.12	92.22	94.08	94.48
Sao Paulo	94.65	91.08	93.26	93.02
Los Angeles	92.96	89.29	91.56	90.18
Shanghai	93.91	90.01	90.34	92.09
Chicago	93.34	90.32	91.31	91.06
Moscow	93.92	90.91	92.87	91.95

however, they failed to extract UGS classes in cloud covered areas, leading to the underestimation of UGS (see Fig. 6 (c)-(e) red box). The proposed network method can solve the underestimated issue because the proposed CAI-net can recognize UGS classes in cloud contaminated areas.

To quantify the performance of the proposed network, four assessment parameters (OA, Kappa, recall, and precision) were calculated for four typical urban scenes. Table 2 show the quantitative results of the four UGS mapping methods. Comparing to DeepLab v3plus and FCN, Geospatial neural network method yielded best values for OA, Kappa, recall, and precision.

From Table 4, the proposed network was found to outperform Geospatial neural network, with average OA of 95.13% versus 81.59%, average Kappa of 91.09% versus 74.91%, average recall of 92.01% versus 76.97%, and average precision of 93.42% versus 77.01%. After the introduction of CAI-net block, the improvement in OA, Kappa, recall, precision was significant. Average OA has increased by 13.54%, and UGS classes in cloud covered areas were correctly extracted. Similarly, average Kappa by 16.18%, average recall by 15.04%, and average precision by 16.41%, as compared to Geospatial neural network. The CAI-net block played an important role for improving UGS mapping in cloud covered areas because it can recognize UGS classes in cloud contaminated areas. The encoder-decoder CAI-net with an attention mechanism can show better performance in UGS mapping from contaminated cloud images.

4.3. *Mapping urban green spaces with different geographical settings*

To further evaluate the mapping performance of the proposed method at different urban scales, we have selected three types of urban patterns (including Beijing, Wuhan, and Guangzhou, China). The different urban patterns have different distributions of UGS due to the composition and fabrics of urban ecosystems (Guo et al., 2020).

Fig. 7 (b1-b3) show the results of extracted urban green spaces by using the proposed network for Beijing, Wuhan, and Guangzhou, China, respectively. By using the cloud detection module, we found that clouds occupied 13.18% in area in Beijing, 15.04% in Wuhan, and in Guangzhou 11.03% in the test Sentinel-2A images. From Fig. 7. (b1-b3), it can be observed that the proposed method can generate spatially continuous UGS maps from the could contaminated Sentinel-2A images in all three types of urban patterns.

Table 5 show quantitative evaluation result using the proposed

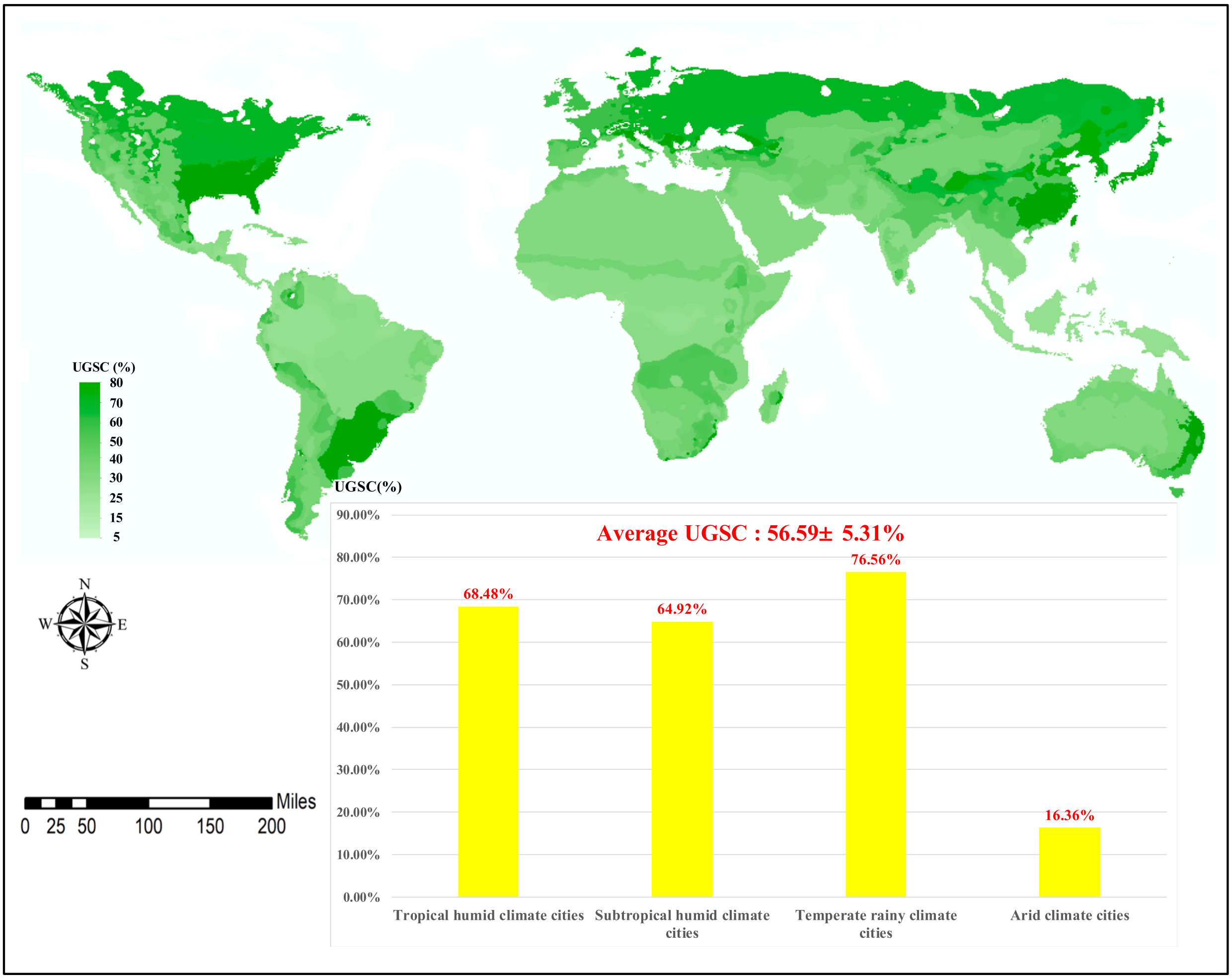

Fig. 11. Urban green space coverage in different climatic zones. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

method in the three cities. Overall mapping results of three types of urban patterns using the proposed method, show an average overall accuracy of 94.38%, an average Kappa of 90.10%, an average recall of 92.43%, an average precision of 93.36%. According to the visualization and quantitation results, the proposed method can be regarded to be effective in generating spatially continuous UGS maps in cloud covered areas.

5. Discussion

5.1. Effect of cloud cover on UGS mapping

Sustainable urban development research is dependent on spatial–temporal continuous UGS data (Venkataramanan et al., 2019). Currently, UGS maps with spatial resolution of 10 m generated from Sentinel-2A images can provide a timely, accurate and essential data for sustainable urban development research and related studies. Cloud cover is inevitable, especially in tropical and subtropical areas, which causes uncertainty in UGS mapping (Zhao et al., 2016). Hence, we discuss the effect of different cloud coverage of Sentinel-2 data on UGS mapping when using the proposed CAI-net method in this section.

To test mapping performance at the different cloud coverage, we have selected 11 images with cloud coverage ranging from 0% to 50%. Table 6 shows details about tested Sentinel-2 images. Fig. 8 shows the effects of the different cloud coverage on UGS mapping accuracy when using the proposed method. From Fig. 8 (a)-(d), we can observe that UGS mapping accuracy of the proposed method with different cloud cover coverage, the OA ranging from 98.36% to 83.71%, the Kappa ranging from 93.56% to 85.56%, the recall ranging from 96.14% to 87.54%, and precision ranging from 93.34% to 85.17%. According to the results, the proposed method can yield a robust mapping accuracy of UGS against different cloud fractions of sentinel-2 images.

5.2. Global UGS mapping in different climatic zones

With increasing urbanization and climate change, global UGS are decreasing at an alarming rate (Sun et al., 2019). According to the United Nations (United Nations, 2019), about 68% of world population will live in urban region by 2050. However, the global UGS dataset does not yet exist by climatic zone. To support urban sustainable development, we chose 1080 cities for producing maps of UGS in different climatic zones. Fig. 9 shows locations of these in different climatic zones. We chose the HLS data and public geographic data as the data sources for mapping global UGS.

The UGS maps of 1080 cities were produced by using the proposed method at global scale. Fig. 10 shows examples of UGS maps in different climatic zones, including Cairo, Sao Paulo, Los Angeles, Shanghai, Chicago, Moscow. From Fig. 10, it can be observed that UGS maps can capture different spatial patterns in global different climatic zones. From Fig. 10 (a)-(b), the spatial pattern of UGS was sparse for Cairo city and Sao Paulo city. From Fig. 10 (c)-(d), we found that the spatial pattern of UGS was compact in Los Angeles and Shanghai. From Fig. 10 (e)-(f), the spatial pattern of UGS was dense in Chicago, Moscow. We observed that UGS was sparse in desert climate cities, whereas the spatial pattern of UGS was dense in tropical climate cities, which could be related to temperature and rainfall of the localities.

Accuracy of the UGS maps was evaluated by examining samples of 500 × 500 pixels as validation regions against manual UGS mapping. Table 7 shows quantitative evaluation results of UGS maps for selected cities in different climate zones. The mapping results of those cities show overall accuracy of 92.96%, Kappa of 89.29%, recall of 90.34%. The proposed method can provide a reliable accuracy mapping of UGS for cities in global different climatic zones.

To explore the spatial patterns of UGS of the cities in different climatic zones, the urban green space coverage (UGSC) indicator was calculated by using the global 1080 cities of UGS maps. In this study, UGSC indicator can be defined as follows:

$$UGSC = \frac{UGS_{pixel}}{Total_{pixel}} \tag{9}$$

where UGS_{pixel} is the number of UGS pixels, and $Total_{pixel}$ is the total number of pixels for an entire urban image.

Fig. 11 shows distribution of urban green space coverage for global different climatic zones. From Fig. 11, the average UGSC was 56.59 ± 5.31% at the global scale. The UGSC of temperate rainy cities yielded more than76.56%, while UGSC of arid climate cities occupied less than 16.36%. In addition, the histogram of global UGSC data indicate that UGSC of temperate rainy cities had higher UGSC values than cities in arid climate zones, tropical humid climate zones, and subtropical humid climate zones.

The main advantages of the proposed method include: (i) The CAI-net was employed to eliminate cloud effects for improving UGS mapping in cloud covered regions; and (ii) The proposed method can generate spatial continuous global UGS data products. This study can provide valuable global UGS data to support the UN SDG 11.

6. Conclusions

Optical satellite derived UGS maps provide a timely, accurate and essential data in sustainable urban development research. However, clouds are a key problem in optical satellite image processing that is highly pertinent to UGS mapping uncertainty. This study exploited how optical satellite data and crowdsourcing geospatial big data can be integrated for generating quality spatially continuous UGS maps. To eliminate clouds effects for UGS mapping, this research proposed an automatic UGS mapping method with the three sequential modules, i.e., auxiliary data pre-processing module, the CAI-net, and non-cloud scenes classification module. The results show an average overall accuracy of 92.14%, average Kappa of 89.12%, average recall of 91.31%, and average precision of 90.21% for 1080 cities in different climatic zones. The main advantages of the proposed automatic UGS mapping method included: (i) Cloud attention intelligent network was designed to eliminate cloud effects to generate spatially continuous UGS maps of different geographical settings; and (ii) The proposed method provided a new insight into the use of AI technology for accurate UGS mapping globally in different climatic zones.

Nevertheless, the proposed CAI-net contains some limitations. First, the CAI-net cannot retrieve detailed UGS classes within clouds when the cloud cover was greater than 50%, because the CAI-net need more auxiliary data to recognize UGS classes in a large of cloud contaminated areas. Secondly, the performance of the CAI-net relied on the quality of auxiliary data because cloud classification module depended on auxiliary data to estimate the probability of UGS classes in cloud covered areas. In future research, it is desired to improve the performance of the proposed CAI-net by integrating multimode data, such as Sentinel SAR data and street view images as auxiliary maps while trying to overcome the above-stated limitations.

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.

Acknowledgment

The authors thank European Space Agency and OpenStreetMap (OSM) platform for providing freely experimental data. This work was supported in part by GuangDong Basic and Applied Basic Research Foundation (Grant No. 2022A1515110890), the National Natural Science Foundation of China (Grant No. 41971405), and the Global STEM Professorship of Hong Kong Special Administrative Region Government.

References

Bolton, D.K., Gray, J.M., Melaas, E.K., et al., 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. Remote Sensing of Environment 240, 111685.

Chaves, M., CA Picoli, M., D. Sanches, I., 2020. Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. Remote Sensing, 12(18), 3062.

Chen, Y., Fan, R., Bilal, M., et al., 2018a. Multilevel cloud detection for high-resolution remote sensing imagery using multiple convolutional neural networks. ISPRS International Journal of Geo-Information 7 (5), 181.

Chen, B., Huang, B., Xu, B., 2017. Multi-source remotely sensed data fusion for improving land cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 124 (2), 27–39.

Chen, W., Huang, H., Dong, J., et al., 2018b. Social functional mapping of urban green space using remote sensing and social sensing data. ISPRS Journal of Photogrammetry and Remote Sensing 146 (12), 436–452.

Chen, J., Liao, A., Cao, X., et al., 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing 103 (5), 7–27.

Chen, Y., Tang, L., Yang, X., et al., 2020. Thick Clouds Removal From Multitemporal ZY-3 Satellite Images Using Deep Learning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13 (2), 143–153.

Chen, Y., Weng, Q., Tang, L., et al., 2021b. Automatic mapping of urban green spaces using a geospatial neural network. GIScience & Remote Sensing 58 (4), 624–642.

Chen, Y., Weng, Q., Tang, L., et al., 2022. Thick Clouds Removing From Multitemporal Landsat Images Using Spatiotemporal Neural Networks. IEEE Transactions on Geoscience and Remote Sensing 60 (1), 1–14.

Chen, B., Xu, B., Gong, P., 2021a. Mapping essential urban land use categories (EULUC) using geospatial big data: Progress, challenges, and opportunities. Big Earth Data, Early Access. https://doi.org/10.1080/20964471.2021.1939243.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., et al., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote sensing of environment 219, 145–161.

Di, S., Li, Z.L., Tang, R., et al., 2019. Urban green space classification and water consumption analysis with remote-sensing technology: a case study in Beijing. China. International Journal of Remote Sensing 40 (6), 1909–1929.

Dong, J.W., Kuang, W.H., Liu, J.Y., 2017. Continuous land cover change monitoring in the remote sensing big data era. Science China Earth Sciences. 60 (12), 2223–2224.

Du, S., Zhang, F., Zhang, X., 2015. Semantic classification of urban buildings combining VHR image and GIS data: An improved random forest approach. ISPRS journal of photogrammetry and remote sensing 105 (7), 107–119.

Foody, G.M., 2020. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. Remote Sensing of Environment 239, 111630.

Fu, X., & Qu, H., 2018. Research on semantic segmentation of high-resolution remote sensing image based on full convolutional neural network. In 2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE) (pp. 1-4). IEEE.

Gong, P., Chen, B., Li, X., et al., 2020. Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. Sci. Bull. 65, 182–187.

Hang, R., Li, Z., Ghamisi, P., et al., 2020. Classification of hyperspectral and LiDAR data using coupled CNNs. IEEE Transactions on Geoscience and Remote Sensing 58 (7), 4939–4950.

Heikinheimo, V., Tenkanen, H., Bergroth, C., et al., 2020. Understanding the use of urban green spaces from user-generated geographic information. Landscape and Urban Planning 201 (9), 103845.

Hong, D., Hu, J., Yao, J., et al., 2021. Multimodal remote sensing benchmark datasets for land cover classification with a shared and specific feature learning model. ISPRS Journal of Photogrammetry and Remote Sensing 178 (8), 68–80.

Huang, C., Yang, J., Clinton, N., et al., 2021. Mapping the maximum extents of urban green spaces in 1039 cities using dense satellite images. Environmental Research Letters 16 (6), 064072.

Jiang, S., Alves, A., Rodrigues, F., et al., 2015. Mining point-of-interest data from social networks for urban land use classification and disaggregation. Computers Environment & Urban Systems 53 (9), 36–46.

Jie, R.A., Yang, S.B., Hw, B., et al., 2021. A two-step mapping of irrigated corn with multi-temporal modis and landsat analysis ready data. ISPRS Journal of Photogrammetry and Remote Sensing 176 (6), 69–82.

Johnson, B.A., Iizuka, K., 2016. Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. Applied Geography 67 (2), 140–149.

Johnson, B.A., Iizuka, K., Bragais, M.A., et al., 2017. Employing crowdsourced geographic data and multi-temporal/multi-sensor satellite imagery to monitor land cover change: A case study in an urbanizing region of the Philippines. Computers, Environment and Urban Systems 64, 184–193.

Kranjčić, N., Medak, D., Župan, R., Rezo, M., et al., 2019. Support Vector Machine Accuracy Assessment for Extracting Green Urban Areas in Towns. Remote Sensing 11 (6), 655.

Lee, A.C.K., Jordan, H.C., Horsley, J., et al., 2015. Value of urban green spaces in promoting healthy living and wellbeing: prospects for planning. Risk Manag. Healthc. Policy 8, 131–137.

Liu, W., Yue, A., Shi, W., et al., 2019. An Automatic Extraction Architecture of Urban Green Space Based on DeepLabv3plus Semantic Segmentation Model. 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC) (pp. 311-315), Taipei, Taiwan, IEEE.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).

Moon, M., Richardson, A.D., Friedl, M.A., 2021. Multiscale assessment of land surface phenology from harmonized Landsat 8 and Sentinel-2, PlanetScope, and PhenoCam imagery. Remote Sensing of Environment 266, 112716.

Nuthammachot, N., Stratoulias, D., 2019. Fusion of Sentinel-1A and Landsat-8 images for improving land use/land cover classification in Songkla province. Thailand. Applied Ecology and Environmental Research 17 (2), 3123–3135.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrology and earth system sciences 11 (5), 1633–1644.

Pekel, J.F., Cottam, A., Gorelick, N., et al., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540 (7633), 418–422.

Peng, H., Li, J., Song, Y., et al., 2017. Incrementally learning the hierarchical softmax function for neural language models. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 31, No. 1).

Preidl, S., Lange, M., Doktor, D., et al., 2020. Introducing APiC for regionalised land cover mapping on the national scale using Sentinel-2A imagery. Remote Sensing of Environment 240, 111673.

Punalekar, S.M., Verhoef, A., Quaife, T.L., et al., 2018. Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model. Remote Sensing of Environment 218 (12), 207–220.

Qu, C., Li, P., Zhang, C., 2021. A spectral index for winter wheat mapping using multi-temporal Landsat NDVI data of key growth stages. ISPRS Journal of Photogrammetry and Remote Sensing 175 (5), 431–447.

Roy, D.P., Huang, H., Boschetti, L., 2019. Landsat-8 and Sentinel-2 burned area mapping-A combined sensor multi-temporal change detection approach. Remote Sensing of Environment 231, 111254.

Salberg, A.B., 2010. Land cover classification of cloud-contaminated multitemporal high-resolution images. IEEE Transactions on Geoscience and Remote Sensing 49 (1), 377–387.

Sharma, A., Liu, X., Yang, X., et al., 2018. Land cover classification from multi-temporal, multi-spectral remotely sensed imagery using patch-based recurrent neural networks. Neural Networks 105 (9), 346–355.

Shi, X., Chen, Z., Wang, H., et al., 2015. Convolutional LSTM network: A machine learning approach for precipitation now casting. In Advances in neural information processing systems 802–810.

Shrestha, D.P., Saepuloh, A., van der Meer, F., 2019. Land cover classification in the tropics, solving the problem of cloud covered areas using topographic parameters. International Journal of Applied Earth Observation and Geoinformation 77 (5), 84–93.

Su, W., Sui, D., Zhang, X., 2020. Satellite image analysis using crowdsourcing data for collaborative mapping: current and opportunities. International Journal of Digital Earth 13 (6), 645–660.

Sukawattanavijit, C., Jie, C., Zhang, H., 2017. Ga-SVM algorithm for improving land-cover classification using sar and optical remote sensing data. IEEE Geoscience and Remote Sensing Letters 14 (3), 284–288.

Sun, C., Lin, T., Zhao, Q., et al., 2019. Spatial pattern of urban green spaces in a long-term compact urbanization process—A case study in China. Ecological indicators 96, 111–119.

UNDESA (Ed.), 2019. World Population Prospects. The 2019 Revision. United Nations Department of Economic and Social Affairs Available online at. https://population.un.org/wpp/ checked on 7/16/2019.

Vaidya, H., Chatterji, T., 2020. SDG 11 Sustainable Cities and Communities. In: Actioning the Global Goals for Local Impact. Springer, Singapore, pp. 173–185.

Vaswani, A., Shazeer, N., Parmar, N., et al., 2017. Attention is all you need. In, Advances in Neural Information Processing Systems. pp. 5998-6008.

Venkataramanan, V., Packman, A.I., Peters, D.R., et al., 2019. A systematic review of the human health and social well-being outcomes of green infrastructure for storm water and flood management. J. Environ. Manage. 246, 868–880.

Voltersen, M., Berger, C., Hese, S., et al., 2014. Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level. Remote Sensing of Environment 154 (11), 192–201.

Weng, Q., 2012. Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. Remote Sensing of Environment, 117, 34-49.

Woo, S., Park, J., Lee, J.Y., Kweon, I.S., 2018. Cbam: Convolutional block attention module. In: In Proceedings of the European conference on computer vision (ECCV), pp. 3–19.

Xie, S., Liu, L., Zhang, X., et al., 2019. Annual land-cover mapping based on multi-temporal cloud-contaminated landsat images. International Journal of Remote Sensing 40 (10), 3855–3877.

Xu, F., Somers, B., 2021. Unmixing-based sentinel-2 downscaling for urban land cover mapping. ISPRS Journal of Photogrammetry and Remote Sensing 171 (1), 133–154.

Zhang, X., Du, S., Wang, Q., 2017. Hierarchical semantic cognition for urban functional zones with VHR satellite images and poi data. ISPRS Journal of Photogrammetry and Remote Sensing 132 (10), 170–184.

Zhang, J., Wang, H., Wang, Y., Zhou, Q., Li, Y., 2021. Deep network based on up and down blocks using wavelet transform and successive multi-scale spatial attention for cloud detection. Remote Sensing of Environment 261 (8), 112483.

Zhao, Y., Feng, D., Yu, L., et al., 2016. Detailed dynamic land cover mapping of Chile: Accuracy improvement by integrating multi-temporal data. Remote Sensing of Environment 183 (9), 170–185.