

## Mapping and analyzing the local climate zones in China's 32 major cities using Landsat imagery based on a novel convolutional neural network

Xin Huang, Anling Liu & Jiayi Li

To cite this article: Xin Huang, Anling Liu & Jiayi Li (2021) Mapping and analyzing the local climate zones in China's 32 major cities using Landsat imagery based on a novel convolutional neural network, Geo-spatial Information Science, 24:4, 528-557, DOI: [10.1080/10095020.2021.1892459](https://doi.org/10.1080/10095020.2021.1892459)

To link to this article: <https://doi.org/10.1080/10095020.2021.1892459>



© 2021 Wuhan University. Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 06 May 2021.



[Submit your article to this journal](#)



Article views: 3047



[View related articles](#)






[View Crossmark data](#)



Citing articles: 20 [View citing articles](#)

# Mapping and analyzing the local climate zones in China's 32 major cities using Landsat imagery based on a novel convolutional neural network

Xin Huang <sup>a,b</sup>, Anling Liu <sup>a</sup> and Jiayi Li <sup>a</sup>

<sup>a</sup>School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China; <sup>b</sup>State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

## ABSTRACT

The Local Climate Zone (LCZ) scheme provides researchers with a standard method to monitor the Urban Heat Island (UHI) effect and conduct temperature studies. How to generate reliable LCZ maps has therefore become a research focus. In recent years, researchers have attempted to use Landsat imagery to delineate LCZs and generate maps worldwide based on the World Urban Database and Access Portal Tools (WUDAPT). However, the mapping results obtained by the WUDAPT method are not satisfactory. In this paper, to generate more accurate LCZ maps, we propose a novel Convolutional Neural Network (CNN) model (namely, LCZ-CNN), which is designed to cope with the issues of LCZ classification using Landsat imagery. Furthermore, in this study, we applied the LCZ-CNN model to generate LCZ mapping results for China's 32 major cities distributed in various climatic zones, achieving a significantly better accuracy than the traditional classification strategies and a satisfactory computational efficiency. The proposed LCZ-CNN model achieved satisfactory classification accuracies in all 32 cities, and the Overall Accuracies (OAs) of more than half of the cities were higher than 80%. We also designed a series of experiments to comprehensively analyze the proposed LCZ-CNN model, with regard to the transferability of the network and the effectiveness of multi-seasonal information. It was found that the first convolutional stage, corresponding to low-level features, shows better transferability than the second and third convolutional stages, which extract high-level and more image- or task-oriented features. It was also confirmed that the multi-seasonal information can improve the accuracy of LCZ classification. The thermal characteristics of the different LCZ classes were also analyzed based on the mapping results for China's 32 major cities, and the experimental results confirmed the close relationship between the LCZ classes and the magnitude of the Land Surface Temperature (LST).

## ARTICLE HISTORY

Received 28 September 2020  
Accepted 15 February 2021

## KEYWORDS

Deep learning; convolutional neural network; local climate zone scheme; land surface temperature

## 1. Introduction

The Urban Heat Island (UHI) effect, i.e. the atmospheric warmth of a city compared with its surrounding countryside, is a significant climatic response to the disruptions caused by urbanization (Stewart and Oke 2012). Measurements of the UHI intensity of the thermal environment have become the key content of recent studies (Streutker 2003; Fisher, Mustard, and Vadeboncoeur 2006; Li et al. 2011). In these studies, most of the researchers have measured the UHI effect through a simple “urban-rural” classification scheme, to show the temperature difference between the urban area and its rural surroundings (Streutker 2003; Fisher, Mustard, and Vadeboncoeur 2006; Li et al. 2011). However, urban-rural systems are typically complex, and the temperature data from various sites can have significant differences due to the distinct thermodynamic characteristics of local landscapes. For these reasons, it is difficult to compare results across cities and to obtain quantitative meta data of the environment around a site using this simple binary system.

In order to address the above problems, the Local Climate Zone (LCZ) scheme was introduced by (Stewart and Oke 2012) as a climate-based classification scheme, and is now regarded as a standard way to monitor the UHI effect and analyze urban temperatures (Xu et al. 2017). Compared with simple urban-rural division, the LCZ system is segmented into “built” types (LCZ 1–10) and “natural” types (LCZ A–G), based on the regional landscape patterns (Stewart and Oke 2012). LCZs can be used to facilitate inter-site comparison and effectively measure the UHI magnitude among different cities (Xu et al. 2017). The classification system can be applied to monitoring the UHI effect, establishing climatic models, and designing urban landscapes.

There have been a number of studies of computer-based approaches to interpret LCZs. One of the widely used techniques is the Geographic Information System (GIS)-based technique (Lelovics et al. 2014; Geletič and Lehnert 2016), in which a large amount of input parameters about the urban surface are needed, such

as the Sky View Factor (SVF), Impervious Surface Fraction (ISF), and building height (Wang et al. 2018). However, detailed urban geographical data are not always available in many cities, especially in the developing countries (Bechtel et al. 2015; Xu et al. 2017). The other common technique is the use of freely available satellite imagery to delineate LCZs and generate maps worldwide. The World Urban Database and Access Portal Tools (WUDAPT) project is a community-based initiative to produce LCZ maps around the world. The WUDAPT method (Ching et al. 2015) involves conducting urban classification based on the LCZ concept with Landsat images, the Google Earth platform, and a random forest classifier. To classify LCZs based on the WUDAPT framework, it is first necessary to create training areas using the Google Earth platform, then collect the spectral features of the Landsat images, and finally achieve LCZ classification via the random forest classifier.

The WUDAPT method has been widely adopted in relevant research (Bechtel et al. 2015; Xu et al. 2017) because the input data and tools can be conveniently and freely employed (Wang et al. 2018). However, although the WUDAPT method has been applied to LCZ classification in more than 50 cities around the world (Danylo et al. 2016; Bechtel et al. 2019), the results for only a few cities have been quantitatively validated to have satisfactory mapping accuracies (Bechtel et al. 2015; Danylo et al. 2016; Wang et al. 2018). For instance, LCZ classification based on the WUDAPT method achieved an OA of 64% in Kyiv in the Ukraine (Danylo et al. 2016), and 62% in Guangzhou in China (Xu et al. 2017). These results indicate that the WUDAPT-based LCZ classification approach is not satisfactory, especially for the dense and highly compact Asian cities, and the use of the approach can significantly affect the subsequent applications (Xu et al. 2017). To obtain more accurate mapping results, researchers have attempted to incorporate other satellite data sources, e.g. the addition of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Synthetic Aperture Radar (SAR) data, which improved the OA by 1–2% and 0.3%, respectively (Xu et al. 2017; Bechtel et al. 2015). Xu et al. (2017) also attempted to take into account Gray-level Co-occurrence Matrix (GLCM) textural features, increasing the OA of the classification results for Wuhan and Guangzhou by 2%. However, to sum up, neither the combined use of other satellite images nor the addition of textural features can substantially strengthen LCZ mapping results. Moreover, the accuracy of the mapping results produced by the WUDAPT method over a large number of other cities still requires further validation (Wang et al. 2018).

LCZs are defined by a combination of surface cover, urban structure, material, and human activities, and

can be regarded as complex scenarios with multiple land-cover types and multiple ground objects, spanning scales of hundreds of meters to several kilometers (Stewart and Oke 2012). Although the addition of textural features did not improve the classification accuracy significantly in the existing literature, these features can actually provide important information on urban structure and can contribute to LCZ interpretation. A possible explanation for the unsatisfactory performance of the WUDAPT method in the current studies is that spectral and textural features are still low-level features, which cannot capture detailed information about the urban landscapes and semantics. In summary, LCZ classification is a very complex problem of remote sensing scene interpretation. For example, the open high-rise (LCZ 4) category is made up of open high-rise buildings, low plants, scattered trees, and soil, and the algorithms based on the WUDAPT method might not be well adapted to such a complicated scene classification task. Therefore, it is necessary to design a more suitable method for LCZ classification from the perspective of scene interpretation, and to substantially improve the quality of the current LCZ mapping products.

Convolutional Neural Networks (CNNs), as a commonly used deep learning method, have the ability to represent higher-level features, and have shown their superiority in semantic understanding (Zhang et al. 2018). CNNs have demonstrated great potential in many remote sensing tasks, e.g. scene classification (Gao et al. 2021), water body extraction (Yu et al. 2018), semantic segmentation (Kemker, Salvaggio, and Kanan 2018), and land-use mapping (Huang et al. 2019). These studies showed that CNN models can substantially outperform the traditional image interpretation methods. Nevertheless, CNN models have rarely been applied to the classification of Landsat images, although they do have the potential to mine the spectral-spatial information.

In general, there are two strategies used to exploit CNN models for remote sensing classification: 1) the use of a pre-trained or fine-tuned CNN; and 2) fully training a CNN from scratch. The first strategy relies on the classical pre-trained CNN networks transferred from an auxiliary domain with natural images. These classical CNN models are directly used as feature extractors, and the number of categories can be modified according to the current sample dataset, and then fine-tuned to conduct the classification. Although some classical pre-trained models have been used in previous studies, e.g. LeNet (Lecun et al. 1998), AlexNet (Krizhevsky, Sutskever, and Hinton 2017), GoogLeNet (Szegedy et al. 2015), and VGGNet (Simonyan and Zisserman 2014), their input channels only refer to one or three spectral channels (RGB). However, Landsat images contain many more spectral bands, i.e. seven multispectral bands with a resolution

of 30 m as well as a panchromatic band with a resolution of 15 m. Therefore, the existing CNN models cannot make full use of the rich spectral information in Landsat images. Moreover, the network structures of the existing CNNs are constructed from high-resolution natural images, and are often configured with a large input window, with a lot of convolutional layers and parameters. The collection of massive training samples from remote sensing images also requires a huge amount of labor, and is very time-consuming (Li, Huang, and Gong 2019). In summary, the aforementioned issues restrict the application of the current CNN models to the mapping of Landsat multispectral images. In contrast, the second strategy, i.e. networks based on training from scratch, allows the network architecture to be customized and parameters set according to the specific requirements, leading to greater flexibility and expandability, especially for Landsat images, which have a relatively low spatial resolution, and for which it is difficult to collect training samples. Some studies have attempted to design and train a small task-specific architecture with the aid of the current remote sensing datasets and the typical filters (Li, Huang, and Gong 2019). These small and fully trained CNNs can avoid the dilemmas of the high cost and risk in training the existing “large-scale” networks with high redundancy and over-parameterization (Li, Huang, and Gong 2019). Therefore, it is interesting and meaningful to build an adaptive CNN model dedicated to the task of LCZ classification using Landsat imagery.

The difference of the thermal properties among the different LCZ categories also deserves comprehensive investigation, from the application viewpoint. However, to date, the relevant studies are inadequate. Alexander and Mills (2014) found that the LCZ D category (low plants) was generally the coolest, while the LCZs with higher impervious and building fractions were found to be warmer in Dublin, Ireland. Geletič and Lehnert (2016) calculated daytime Land Surface Temperature (LST) differences for the LCZs in Prague and Brno in the Czech Republic. The highest LST was found in LCZ 2 (compact mid-rise), LCZ 3 (compact low-rise), and LCZ 10 (heavy industry), while the lowest LST was observed in LCZ A (dense trees) and LCZ G (water) in both cities. Lelovics et al. (2014) found that the high-rise and mid-rise LCZs were warmer than the low-rise types in Szeged in Hungary, while Wang et al. (2018) obtained a different result, in that the low-rise and mid-rise LCZs exhibited higher LST than the high-rise types in Phoenix and Las Vegas in the U.S. Most of the recent research (Geletič and Lehnert 2016; Cai et al. 2018; Lelovics et al. 2014) has been aimed at typical cases of one to two cities, and their conclusions about the thermal properties of the different LCZs have not been entirely consistent, owing to the differences of

the urban landscapes and background climates. Although some studies obtained analysis results in multiple cities, they were either based on urban agglomeration from a specific climatic zone (Cai et al. 2018) or a separate analysis of each city (Bechtel et al. 2019), and they did not comprehensively analyze the characteristics and differences of the climatic zones. Stewart, Oke, and Kravayehoff (2014) encouraged more LCZ evaluation studies for urban environments under various climatic conditions. It is therefore necessary to explore the relationship between LCZs and the thermal environment, and to evaluate the LCZ classification scheme performance based on a large number of cities distributed in different climatic zones.

In this context, from the technological aspect, we propose a novel CNN framework (namely, LCZ-CNN), which is designed to cope with the issues of LCZ classification using Landsat imagery. The proposed LCZ-CNN model adequately considers the multi-spectral and multi-seasonal information of the Landsat images. Meanwhile, the traditional classification strategies, including WUDAPT (Ching et al. 2015), the Bag of Visual Words (BoVW) model (Yang and Newsam 2010), and the Multi-Layer Perceptron (MLP) (Hu 2011), were also applied for comparison. Furthermore, we applied the LCZ-CNN model to generate the LCZ mapping results for China's 32 major cities distributed in different climatic zones, and the thermal characteristics of the different LCZ classes were then analyzed.

## 2. Study area and data

### 2.1. Study area

China is a country with a vast territory and a huge population, located in the east of Asia and the west coast of the Pacific Ocean. China's terrain is high in the west and low in the east. Mountains, plateaus, and hills account for about 67% of the land area, while basins and plains account for the rest. Owing to the tremendous differences in geographical conditions, the climate of China is diverse, ranging from tropical in the south to subarctic in the north, and alpine in the higher elevations of the Tibetan Plateau. The diversity of climate and the uneven distribution of population also lead to large differences in the landscapes of the various cities in China.

As the largest developing country in the world, China has experienced a rapid urbanization process in recent decades, exacerbating the UHI effects in the large cities (Gong et al. 2012). Hence, this study focused on the 32 major cities of China, with regard to their accurate LCZ mapping, as well as a thermal environmental analysis (Figure 1). All of these cities are municipality or provincial capitals, which are

located in different climatic zones and exhibit complex urban forms and landscapes. Therefore, the LCZ mapping in these selected cities was both challenging and meaningful, allowing us to not only test the robustness of the proposed deep neural network from the methodology point of view, but also to investigate the thermal characteristics of the megacities of China.

Figure 2 shows examples of the representative LCZ classes in the 32 major cities, with detailed descriptions, as well as the high-resolution images obtained from Google Earth. The numbers of LCZ classes in the different cities are different, due to the differences of their development levels and urban landscapes. In this research, the size of the input window for the LCZ classification was set to 240 m, which was identified

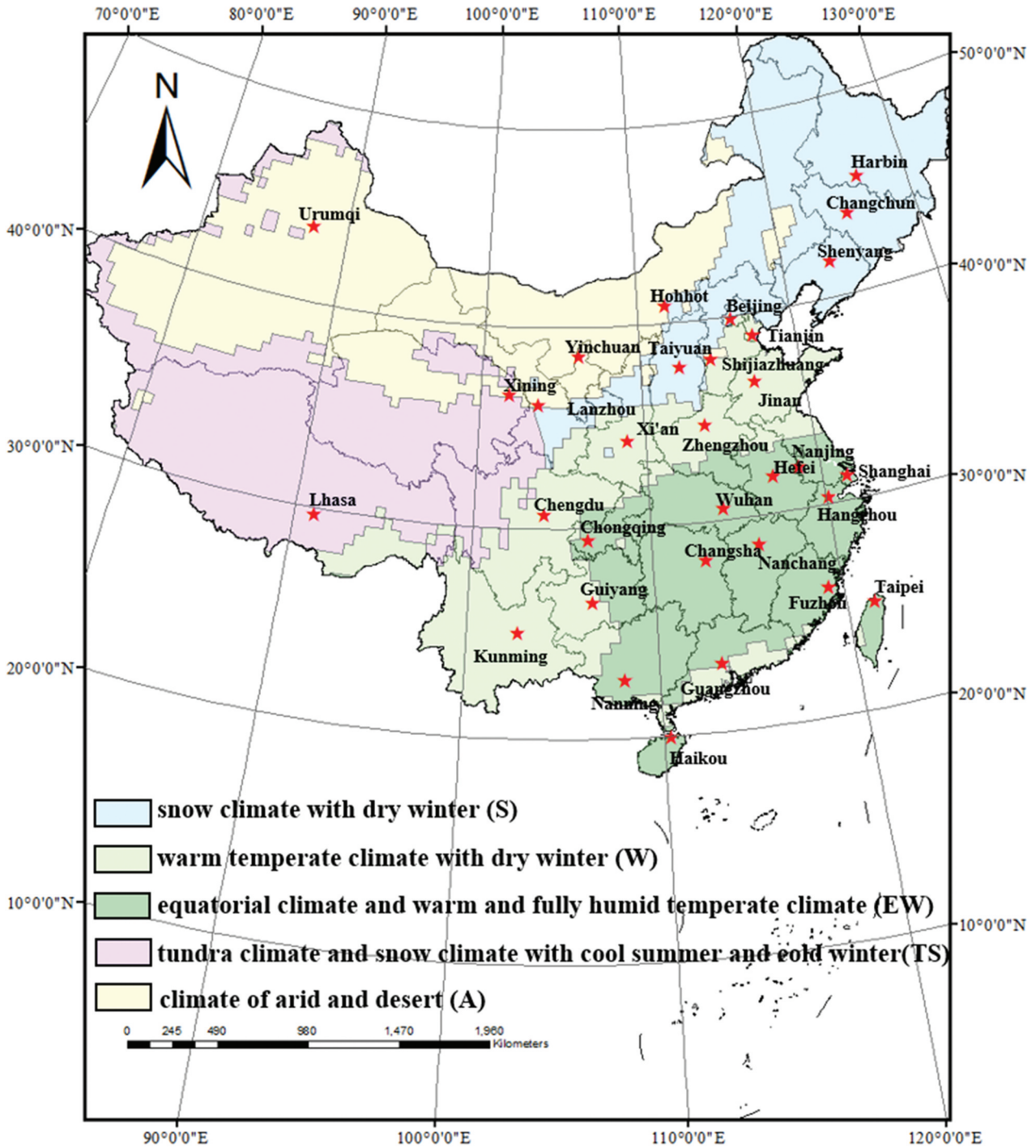
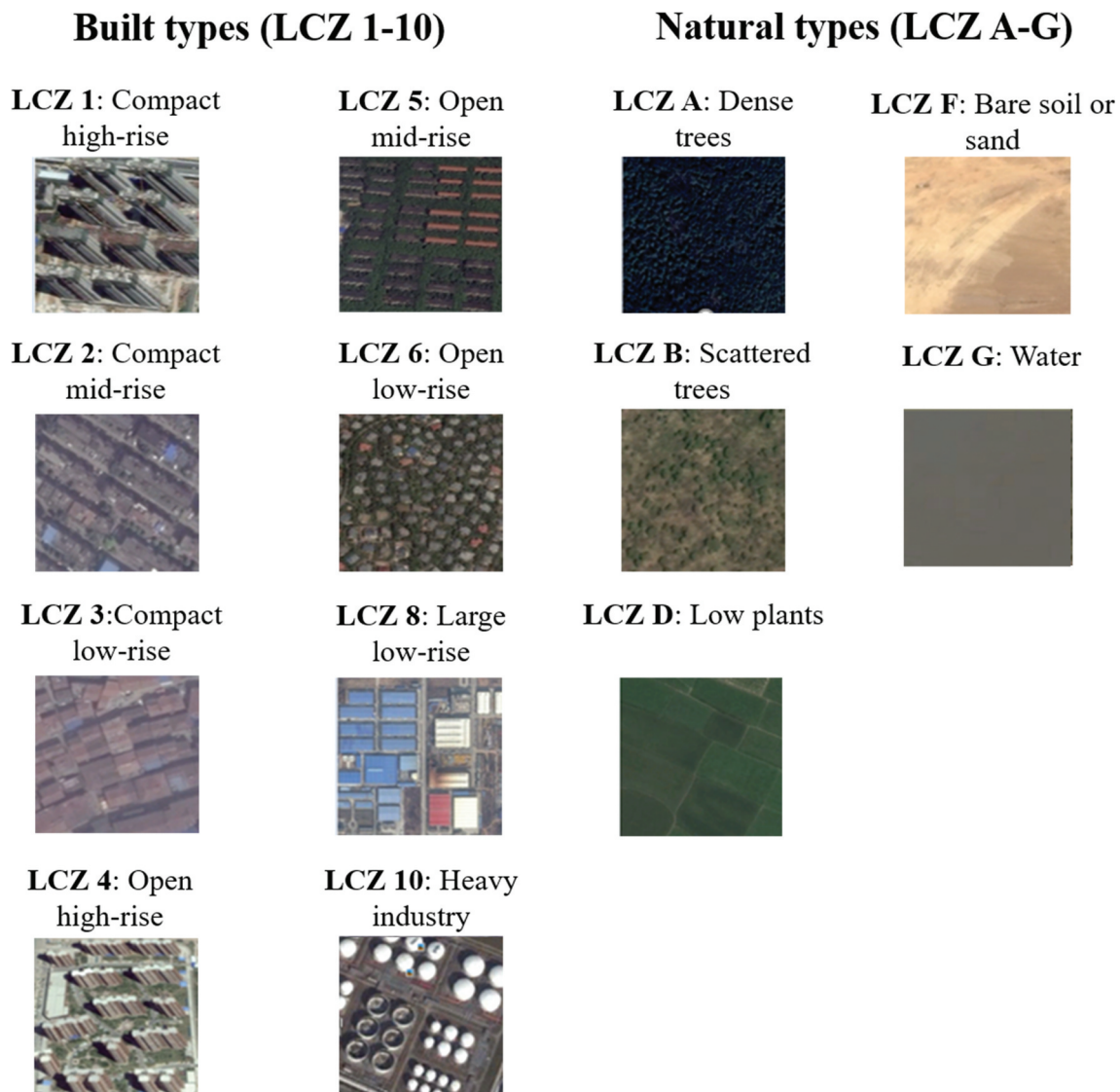


Figure 1. Locations of the 32 major cities and five climatic zones in China. The climatic zones are based on Yang, Huang, and Li (2017).



**Figure 2.** Examples of representative LCZs from Chinese cities. This classification system is segmented into “built” types (LCZ 1–10) and “natural” types (LCZ A–G), based on the regional landscape patterns (Stewart and Oke 2012).

according to the definition of LCZs and the previous experimental results obtained for high-density Asian cities (Stewart and Oke 2012; Lau et al. 2015).

## 2.2. Data

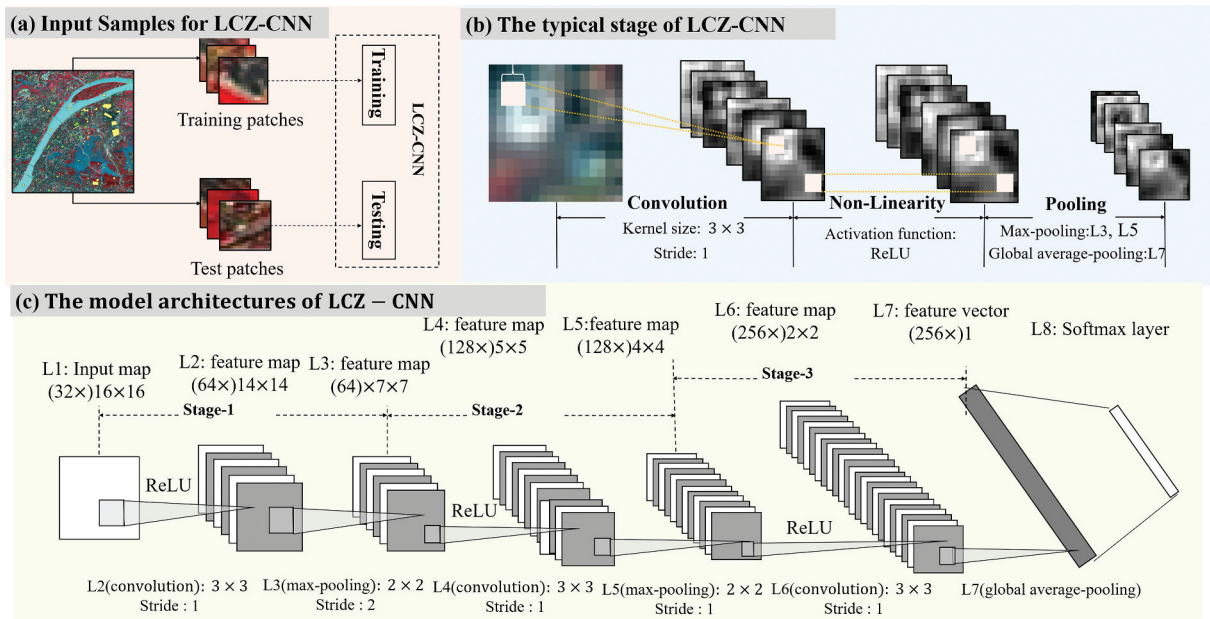
In total, 128 Landsat 8 level-1 image scenes from different seasons acquired mainly in 2015 with minimum cloud coverage were downloaded from the U.S. Geological Survey Earth Explorer website for the 32 cities (Figure A1). Since the image quality for several cities in 2015 was poor, the available images of the most recent years were selected as supplements. We adopted the Additive Wavelet Luminance Proportional (AWLP) method to incorporate the spatial details in the high-resolution (i.e. 15 m) panchromatic bands of the Landsat images, allowing us to better capture the spatial structure information of the LCZs. AWLP is a widely used and robust pan-sharpening algorithm (as suggested by the large-scale comparison in (Meng et al.

2020)), which utilizes the electromagnetic spectrum responses of the sensors to improve the spatial resolution of multispectral images (Otazu et al. 2005).

## 3. Method

### 3.1. Architecture of the proposed LCZ-CNN model

In this research, we developed the novel LCZ-CNN model, adapting deep learning to Landsat-based LCZ classification. The model architecture is shown in Figure 3. The proposed LCZ-CNN model consists of one input layer, three convolutional layers, two max-pooling layers, one global average pooling layer, and one softmax layer (Figure 3(c)). A stage of the LCZ-CNN model (Figure 3(b)) includes a convolution operation, an elementwise non-linear function, and a pooling operation. As such, the main structure of the LCZ-CNN model can be summarized into three typical convolutional stages.



**Figure 3.** The LCZ-CNN framework used in this work. (a) Production process of training dataset and a test dataset. (b) A stage of the LCZ-CNN model. (c) Architecture of the proposed LCZ-CNN model.

The input layer is intended to integrate Landsat images from four seasons to provide temporal information that facilitates LCZ classification (Wicki and Parlow 2017). The input window size ( $16 \times 16$ ) of a sample patch is determined according to the scale ( $240 \text{ m} \times 240 \text{ m}$ ) of the LCZs, which is set according to the scale of the LCZ classification of high-density cities (Lau et al. 2015). The experimental results based on the proposed method in Table A1 also confirm that  $240 \text{ m}$  is a suitable window size for LCZ classification. The number of input channels is 32, consisting of eight Landsat spectral bands with four seasons. A major advantage of the proposed LCZ-CNN model is that it can make full use of the rich temporal and spectral information embedded in the Landsat images. The sample dataset was further divided into a training dataset and a test dataset (Figure 3(a)). The former was used to construct the network, and the latter was used to evaluate the mapping accuracy.

The purpose of the convolutional layers is to extract feature maps by applying a convolution operation to the input bands. Each convolution neuron processes data only for its receptive field. An elementwise non-linear activation function is then applied to these feature maps to complete a non-linear transform. The operation of a convolutional layer can be expressed as:

$$X^l = f(w^l X^{l-1} + b^l) \quad (1)$$

where  $X^l$  denotes the activation value of layer  $l$ ,  $X^{l-1}$  represents the input activation value produced by layer  $(l-1)$ ,  $f(\cdot)$  corresponds to the activation function, and  $w^l$  and  $b^l$  represent the weights and biases at layer  $l$ , respectively.

The Rectified Linear Unit (ReLU) is chosen as the non-linear activation function, due to its ability to alleviate the problem of gradient disappearance, and to speed up the training (Zhang et al. 2016):

$$\text{ReLU}(x) = \max(0, x) \quad (2)$$

where  $x$  represents the output values for the neurons.

The first convolutional layer has 64 small convolutional filters (each of  $3 \times 3$ ). Its function is not only to extract the low-level visual features (Zhang, He, and Lu 2019), but also to output the linear combination of the spectral bands from the multi-seasonal images to sufficiently exploit the spatio-temporal information from the Landsat data. Please note that, in the proposed LCZ-CNN model, the size of an LCZ patch ( $16 \times 16$ ) is much smaller than that for natural images, e.g.  $224 \times 224$  in GoogLeNet (Szegedy et al. 2015) and VGGNet (Simonyan and Zisserman 2014). With regard to the coarse spatial resolution of Landsat imagery ( $15 \text{ m}$  after pan-sharpening), large convolutional filters cannot capture the small spatial variation within hundreds of meters. Thus, small convolutional filters ( $3 \times 3$ ) are also adopted in the fourth and sixth layers, which also make more non-linear changes (i.e. by activation function), as well as reducing the parameters.

The purpose of the pooling layers is to achieve translation invariance, making the architecture less susceptible to small variations in location, and allowing the network to represent the more abstract features of the input as the network depth increases (Lai et al. 2019). The operation of a pooling layer can be written as:

$$N = \frac{W - F + 2 \times P}{S} + 1 \quad (3)$$

where  $N$  and  $W$  denote the sizes of the output maps and input maps, respectively,  $F$  represents the size of the pooling filters.  $P$  corresponds to the number of padding pixels, and  $S$  represents the stride value.

Feature maps obtained from the last convolutional layer are usually fed into the fully connected layers (Figure 4(a)), followed by a softmax classification layer with the same number of classification categories. However, the redundancy of the parameters produced by the fully connected layers can result in overfitting, thus restricting the generalization ability of the whole network (Xu et al. 2019). Moreover, the previously extracted features cannot maintain their spatial structure after passing through the fully connected layers. Therefore, in the LCZ-CNN model, we adopt the strategy of Global Average Pooling (GAP) to replace the fully connected layers (Hsiao et al. 2019). The idea of GAP is to compute the average value of each feature map and input the result vector directly into the soft-max layer (Figure 4(b)). There is no parameter to tune in the GAP layer, so that overfitting can be avoided, and the spatial information is summed to make the spatial translations of the input more robust.

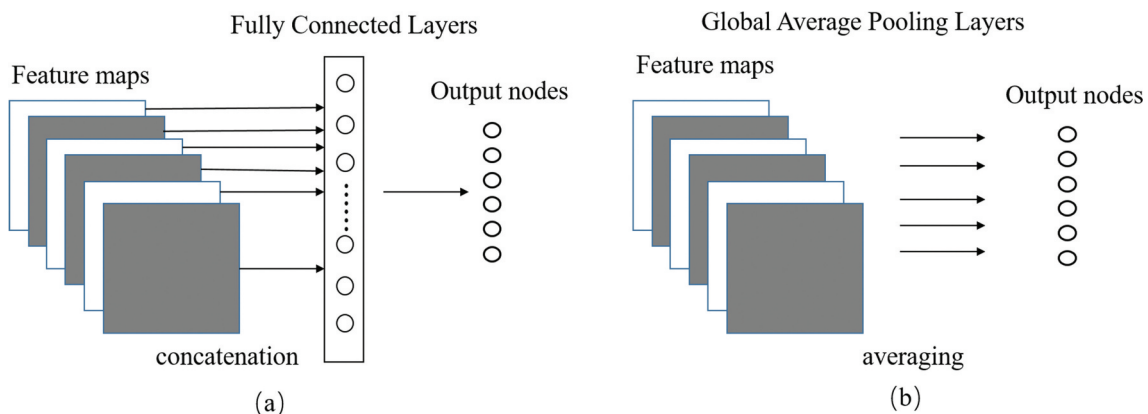
To summarize, the LCZ-CNN model is a lightweight network, specially designed for the task of LCZ mapping using multi-spectral and multi-seasonal Landsat images. The input window size is  $16 \times 16$ , and there are only three convolutional layers in this framework. The employment of the GAP layer also avoids the parameter redundancy caused by the fully connected layers. The proposed lightweight network can obtain satisfactory results with only a moderate amount of training samples.

### 3.2. LCZ-CNN model training

In the training process, we adopted two strategies – dropout and batch normalization – to prevent overfitting. Batch Normalization (BN) is an algorithm for accelerating the convergence and stability of neural networks (Li et al. 2018). When training the LCZ-CNN model, we added a BN operation after each convolutional layer. This can prevent the training from getting stuck in the “saturated regimes of nonlinearities” by normalizing the activation of the whole network (Li et al. 2018). In addition, a batch-normalized architecture allows for faster learning rates, thus generating models with better generalization capabilities.

Dropout is a strategy used to reduce overfitting and to prevent complex co-adaptations on the training data (Srivastava et al. 2014). The specific operation of this method is to set the output of some hidden neurons to zero, which means that each hidden neuron is randomly omitted from the network, with a probability of  $p$ . Thus, the dropped neurons do not participate in the forward-pass and are not used in the back-propagation process. In the different training epochs, different neural networks are formed by dropping neurons randomly. Finally, the neural networks produced in the different training epochs are approximately combined. Therefore, the dropout method prevents overfitting, and the neurons can learn more representative features (Srivastava et al. 2014). In this study, the probability of dropout was set to 0.5, at which point the most random network structures are generated (Srivastava et al. 2014).

Stochastic Gradient Descent (SGD) was selected as the optimization algorithm for LCZ-CNN in the training process. In the phase of model training, we randomly selected 70% of the training data as the training



**Figure 4.** Comparison between a fully connected layer and global average pooling. (a) Operation of fully connected layers. (b) Operation of global average pooling layers.



samples, and the remaining 30% were used as the validation samples. The training data were collected from the Google Earth high-resolution images of the same period. Full details of the production process can be found at <http://www.wudapt.org/>.

### 3.3. Accuracy assessment

In the phase of inference, we used completely independent ground-truth data to evaluate the test accuracy of the different models. The ground-truth production process was the same as the training sample collection process, but was conducted independently. Table A2 shows the training and test data for the 32 cities. The confusion matrix was then calculated from the test samples for the accuracy assessment. Five quality indices calculated from the confusion matrix were then adopted to assess the classification performance (Foody 2002; Stehman 2013), in terms of the Overall Accuracy (OA), Kappa, User's Accuracy (UA), Producer's Accuracy (PA), and F1-score (F).

### 3.4. Analysis of the thermal properties of the LCZ classes

The previous studies (Geletič and Lehnert 2016; Lelovics et al. 2014; Wang et al. 2018) have usually been conducted on a small number of cities (usually

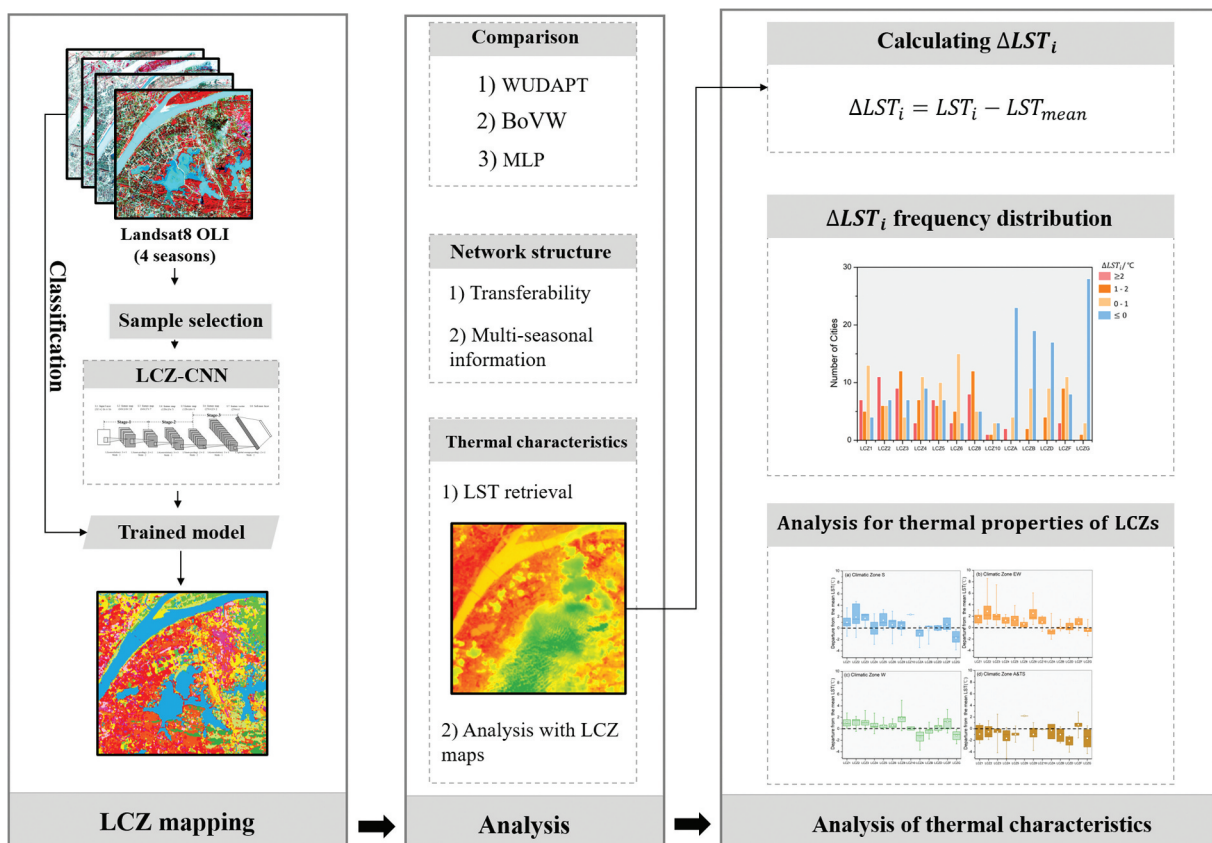
one to two cities), and their conclusions about the relationship between LCZs and LST have not been entirely consistent, owing to the different urban landscapes and background climates. Thus, in this study, we further investigated the thermal characteristics of the LCZ classes in the summer daytime for China's 32 cities, based on our mapping results. The LST data were derived from the thermal infrared images obtained from the Landsat 8 Thermal Infrared Sensor (TIRS). Please note that the thermal infrared bands of the Landsat 8 TIRS were not considered in the previous LCZ classification since they would be used in the subsequent LST analysis. To estimate the LST, we adopted the method of Li et al. (2011), which is a single-window algorithm based on Band 10 of the Landsat 8 TIRS.

To explore the relationship between LCZs and the thermal environment, we calculated the departure from the mean LST for all the LCZ classes:

$$\Delta LST_i = LST_i - LST_{mean} \quad (4)$$

where  $i$  denotes an LCZ class,  $\Delta LST_i$  is the departure from the mean LST for the  $i_{th}$  LCZ class,  $LST_i$  is the mean LST for the  $i_{th}$  LCZ class, and  $LST_{mean}$  is the average LST of all the LCZ classes.

We summarized the  $\Delta LST_i$  frequency distribution in the 32 cities to evaluate the heat/cold island effects for each LCZ class. Box charts of the  $\Delta LST_i$  of the cities



**Figure 5.** The overall processing and analysis flowchart of this research, which is mainly divided into LCZ mapping, analysis of mapping results and analysis of thermal characteristics.

in the different climatic zones were calculated separately, to analyze the LST pattern of the LCZs. Figure 5 shows the entire mapping and analysis flowchart of this research.

## 4. Results

### 4.1. LCZ classification of China's 32 major cities

In this section, we describe how the methods of WUDAPT (Ching et al. 2015), BoVW (Yang and Newsam 2010), and MLP (Hu 2011) for LCZ classification with Landsat imagery were compared with the proposed LCZ-CNN model.

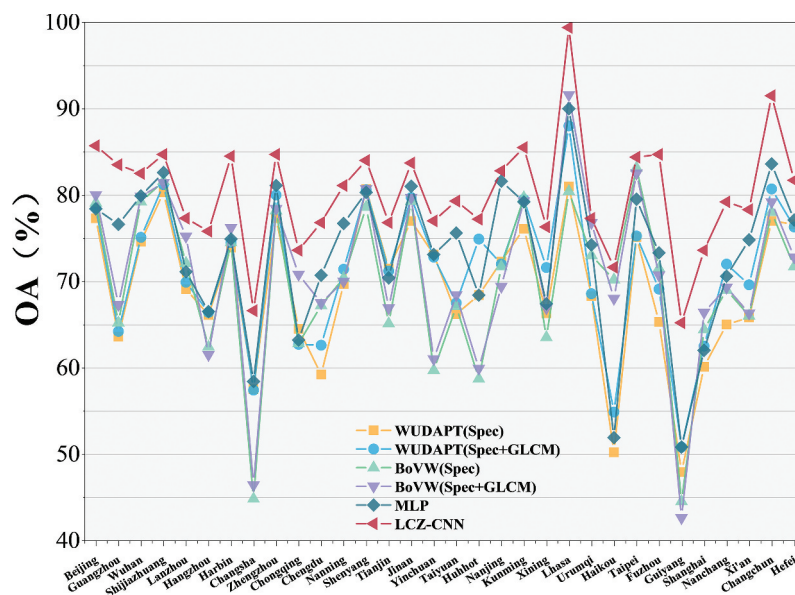
The WUDAPT method was implemented with two different input features, using only the “spectral” features (i.e. WUDAPT (Spec)) and using both the “spectral” and “textural” features (i.e. WUDAPT (Spec+GLCM)), respectively. The spectral features included the eight Landsat bands in four seasons, which was consistent with the input of the LCZ-CNN model. The textural features included four GLCM features, i.e. contrast, homogeneity, energy, and correlation. The window sizes used for calculating the spectral features (average value and standard deviation) and GLCM features were  $16 \times 16$  and  $17 \times 17$ , respectively.

The BoVW model was also implemented with two different input features, using only the “spectral” feature (i.e. BoVW (Spec)) and both the “spectral” and “textural” features (i.e. BoVW (Spec+GLCM)) to create the visual dictionaries (Yang and Newsam 2010). Considering the classification accuracy and the computational efficiency, the size of the visual words for the “spectral” dictionary and “textural” dictionary was set to 128 and 16, respectively, by manual parameter tuning. Random forest was selected as the classifier for

WUDAPT and BoVW, as suggested by the other LCZ classification experiments (Bechtel et al. 2015; Xu et al. 2019), and its robustness to the training set size and noise. The parameters of the MLP were also tuned manually. The learning rate was chosen as 0.001, the number of hidden layers was set to 3, and the number of nodes in each hidden layer was set to {300,200,100}.

The OAs of the experiments are shown in Figure 6. The proposed LCZ-CNN model achieves satisfactory classification accuracies in all 32 cities, ranging from 65.2% to 99.4%. In fact, the OA for more than half of the cities is higher than 80%. The best results are achieved for Kunming, Beijing, Changchun, and Lhasa, with OA values of higher than 85%. The worst results are obtained for Guiyang and Changsha, with OA values of 65.2% and 66.6%, respectively. The increase in OA is very significant when compared with the other benchmark methods, and the proposed LCZ-CNN model achieves the best classification performance in all 32 cities. In more than 20 cities, the OA values of the LCZ classification obtained by WUDAPT are less than 80%, indicating that the LCZ mapping results for some cities (especially the dense and highly compact Asian cities) obtained with the traditional WUDAPT method are not satisfactory, and may affect the subsequent applications (Xu et al. 2017). Both BoVW and MLP achieve better OA values than WUDAPT, but the improvements are not as significant as for LCZ-CNN.

Based on all the test samples from the 32 cities, a confusion matrix was calculated for each method, which are presented in Tables A3–A8. Although considerable confusion among the built types (LCZ 1–10) exists, the confusion between LCZ 1 (compact high-rise) and LCZ 4 (open high-rise), and LCZ 2 (compact mid-rise) and LCZ 5 (open mid-rise), for example, can



**Figure 6.** Overall accuracies for China's 32 major cities. The proposed LCZ-CNN model was compared with other benchmark methods (WUDAPT, BoVW, and MLP).

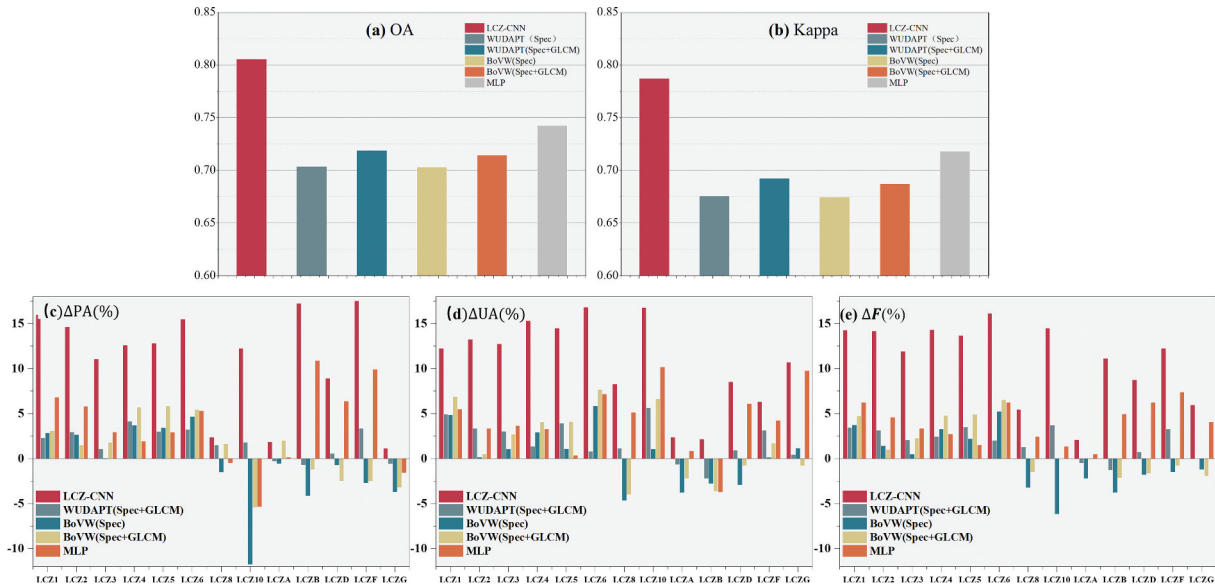


Figure 7. Comprehensive comparison of the classification accuracies of the different methods.

be effectively alleviated by the proposed LCZ-CNN model. Figure 7 shows the OA and Kappa values of the various methods, and the  $\Delta PA$ ,  $\Delta UA$ , and  $\Delta F$  values for each LCZ class. To compare the results, we calculated the differences of the per-class accuracies (PA, UA, and F1-score) between each method and WUDAPT (Spec), which are recorded as  $\Delta PA$ ,  $\Delta UA$ , and  $\Delta F$ , respectively. It can be seen that the advantage of the proposed LCZ-CNN model is obvious, while the other benchmark methods (e.g. WUDAPT (Spec + GLCM), BoVW (Spec), BoVW (Spec + GLCM), and MLP) do not achieve significant accuracy improvements. The superiority of the proposed LCZ-CNN model is further demonstrated by the per-class mapping accuracies (Figure 7(c–e)), where the  $\Delta PA$ ,  $\Delta UA$ , and  $\Delta F$  values of all the LCZ classes are greatly increased compared with WUDAPT (Spec). For most built types (LCZ

1–10), the per-class mapping accuracies of WUDAPT (Spec+GLCM), BoVW (Spec), and BoVW (Spec + GLCM) are also improved, compared with WUDAPT (Spec), which proves that adding texture information and extracting spatial features through the BoVW model can help to improve the classification accuracies for complex scenarios consisting of multiple land-cover types. In addition, it is also interesting that BoVW (Spec) improves the per-class mapping accuracies of most built types, compared with WUDAPT (Spec), and the same situation is also apparent when comparing BoVW (Spec+GLCM) with WUDAPT (Spec+GLCM). This phenomenon indicates that mapping spectral and texture features into sparse mid-level features through the BoVW model can also help to improve the classification accuracy for complex urban scenes. The classification results obtained by MLP are inferior to the

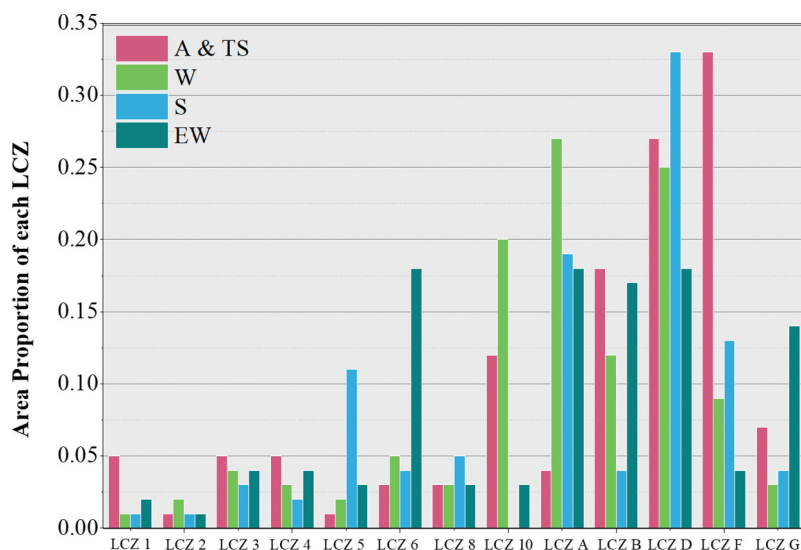


Figure 8. Area proportion of each LCZ in the different climatic zones.

results of LCZ-CNN, overall, revealing the great potential of machine learning algorithms in Landsat image scene classification.

#### 4.2. LCZ mapping results for China's 32 major cities

To explore the relationship between LCZ distribution and climatic zones, the average area proportion of each LCZ class in the different climatic zones is shown in Figure 8. There is only one city (Lhasa) in climatic zone TS (tundra climate and snow climate with cool summer and cold winter). Since climatic zone A (climate of arid steppe and desert) and TS are both located in Northwest China, the cities in these two climatic zones are merged in the subsequent analysis.

Concerning climatic zones A and TS, the natural type with the highest area proportion is LCZ F (bare soil or sand), since the arid climate is detrimental to vegetation growth, and there are many deserts around these cities (Hu, Su, and Zhang 1988). Compared with the other climatic zones, the area proportion of LCZ 1 (compact high-rise buildings) in climatic zones A and TS is also significantly higher. The surrounding topographic conditions (e.g. surrounding deserts, narrow valleys, mountains) and harsh climatic conditions (e.g. scarce precipitation, hot weather) limit the speed and scale of urban expansion in these zones, and hence the buildings are usually compact and high (Roeser et al. 2012).

As for the cities in climatic zone W (warm temperature climate with dry winter), LCZ 10 (heavy industry) and LCZ A (dense trees) occupy a larger area than in the other climatic zones. Most of the cities in climatic zone W (e.g. Tianjin, Taiyuan, Chengdu, Shijiazhuang, and Xi'an) are China's important industrial cities, with prosperous industrial construction. Furthermore, the mild climate in this zone is suitable for plant growth.

LCZ 5 (open mid-rise) is the predominant built type in climatic zone S (snow climate with dry winter), and it occupies a larger proportion than in the other climatic zones, while the area proportion of LCZ 6 (open low-rise) in climatic zone EW (equatorial climate and warm and fully humid temperate climate) is notably high. Most cities in climatic zones S and EW are located on plains or basins with agriculture, and hence there is a lot of space to satisfy the horizontal urban sprawl, forming a lot of open mid- or low-rise types. Moreover, most of the cities in climatic zone EW are located in southern China, with abundant precipitation and large lakes or rivers, and the area of LCZ G (water) in this zone is generally much larger than that in the other climatic zones.

Figure 9 presents the mapping results for four representative cities in different climatic zones. To

show more details of the LCZ distribution, zoomed-in regions of 5 km  $\times$  5 km for each result are also shown in Figure 9, with the original Landsat images and the results obtained by the proposed LCZ-CNN model. It can be observed that the maps are consistent with the characteristics shown in Figure 8. In general, LCZ 1 (compact high-rise), LCZ 2 (compact mid-rise), LCZ 4 (open high-rise), and LCZ 5 (open mid-rise) are the predominant built types in the downtown areas, most of which belong to residential and commercial areas. LCZ 8 (large low-rise) and LCZ 10 (heavy industry) are generally located in suburbs or rural areas. In summary, the LCZ categories and the urban structures in the complicated and high-density Chinese megacities are effectively delineated by the proposed LCZ-CNN model.

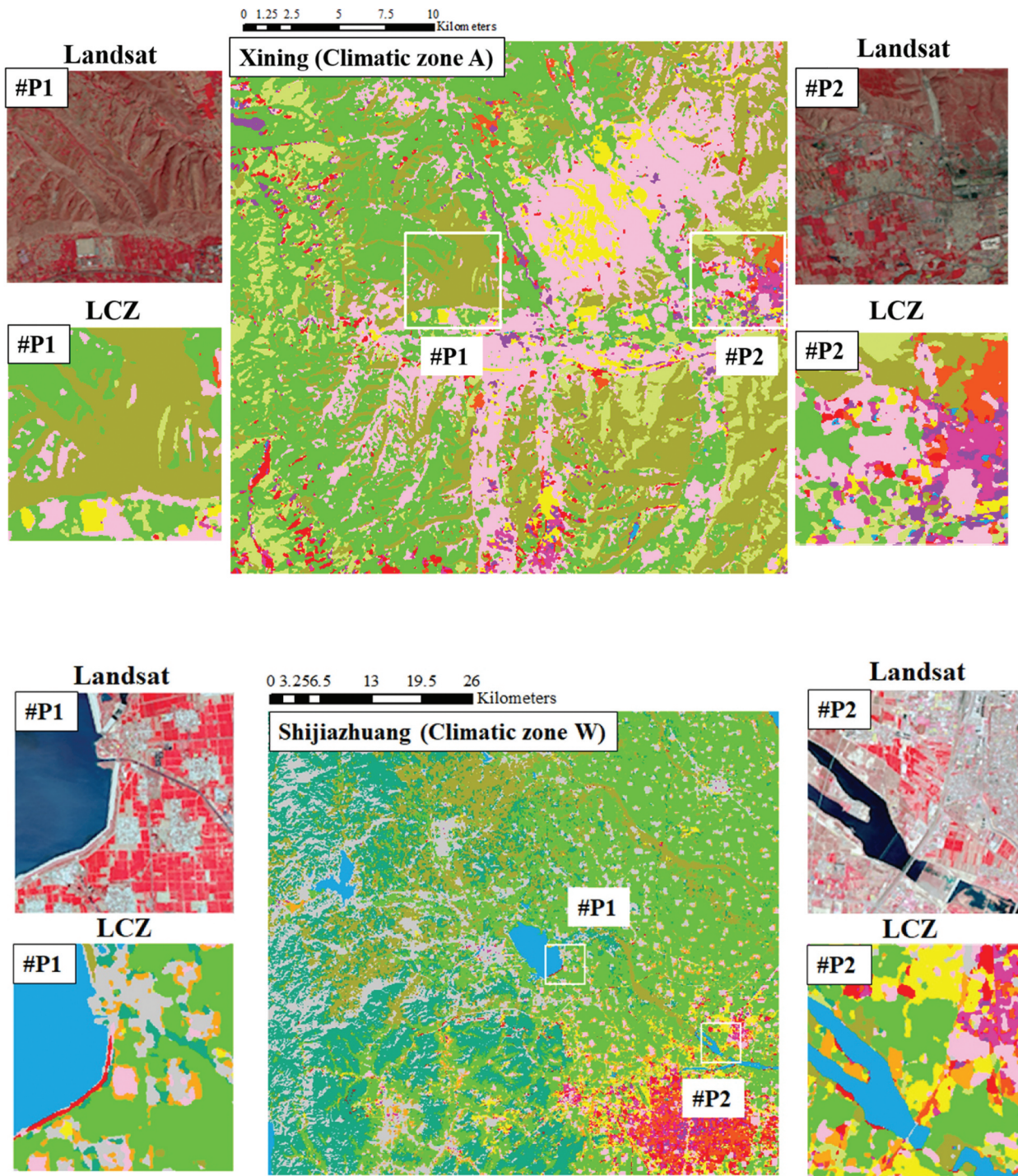
## 5. Discussion

### 5.1. Comparison with other benchmark methods

In the previous sections, we compared the proposed LCZ-CNN model with other benchmark methods (WUDAPT, BoVW, and MLP) in terms of classification accuracy. In this section, we further compare these methods from the perspective of theory and computational efficiency.

The LCZ built types usually correspond to complex urban scenes with man-made structures and other spectrally heterogeneous features. Therefore, the results obtained by WUDAPT (Spec) based only on the image spectral features are not satisfactory, as expected, since adequate spatial information is not utilized. The GLCM texture counts the frequency distribution, which describes how often two gray levels occur in a given spatial relationship, capturing the spatial relationships and contextual information in the neighborhood of the current pixel (Huang et al. 2020). In this way, the complex scenes composed of multiple land-cover types in the LCZ categories can be more effectively represented by the addition of the spatial context relations, which explains why the classification results of WUDAPT (Spec + GLCM) are better than those of WUDAPT (Spec). The experimental results obtained by (Xu et al. 2017) in Wuhan and Guangzhou in China also verified that the quality of LCZ mapping results can be improved with the addition of textural information.

The BoVW model can extract mid-level semantic features of complex scenes from the low-level visual features such as the spectra and texture (Aslam et al. 2019), and is better able to mine the spatial information in remote sensing images. As a result, the classification accuracies for the representative built types (e.g. LCZ 1, 4, 5, and 6) obtained by the BoVW model show significant improvements when compared with



**Figure 9.** Mapping results for four representative cities (Xining, Shijiazhuang, Changchun, and Nanjing).

the results of the WUDAPT method. Please note that the BoVW model is a scene classification method, which has been successfully applied to land-use scene classification (Muhammad et al. 2019) and image retrieval (Arun, Govindan, and Kumar 2020) from high-resolution remote sensing images. However, to date, few studies have applied the BoVW model to information extraction from Landsat imagery. Our experimental results indicate that the BoVW model can achieve a satisfactory accuracy for Landsat image scene classification. However,

the BoVW model calculates the histogram of all the visual words in an image scene, and thus it runs the risk of losing the spatial arrangement information (Yang and Newsam 2010). In addition, the model construction and feature mapping of BoVW is actually an unsupervised process, which may lead to a reduction in the classification accuracy.

The MLP is one of the most popular neural network models in remote sensing thematic mapping. It has been widely used in land-cover mapping (Jamali 2020), land-use classification (Subiyanto et al. 2019),

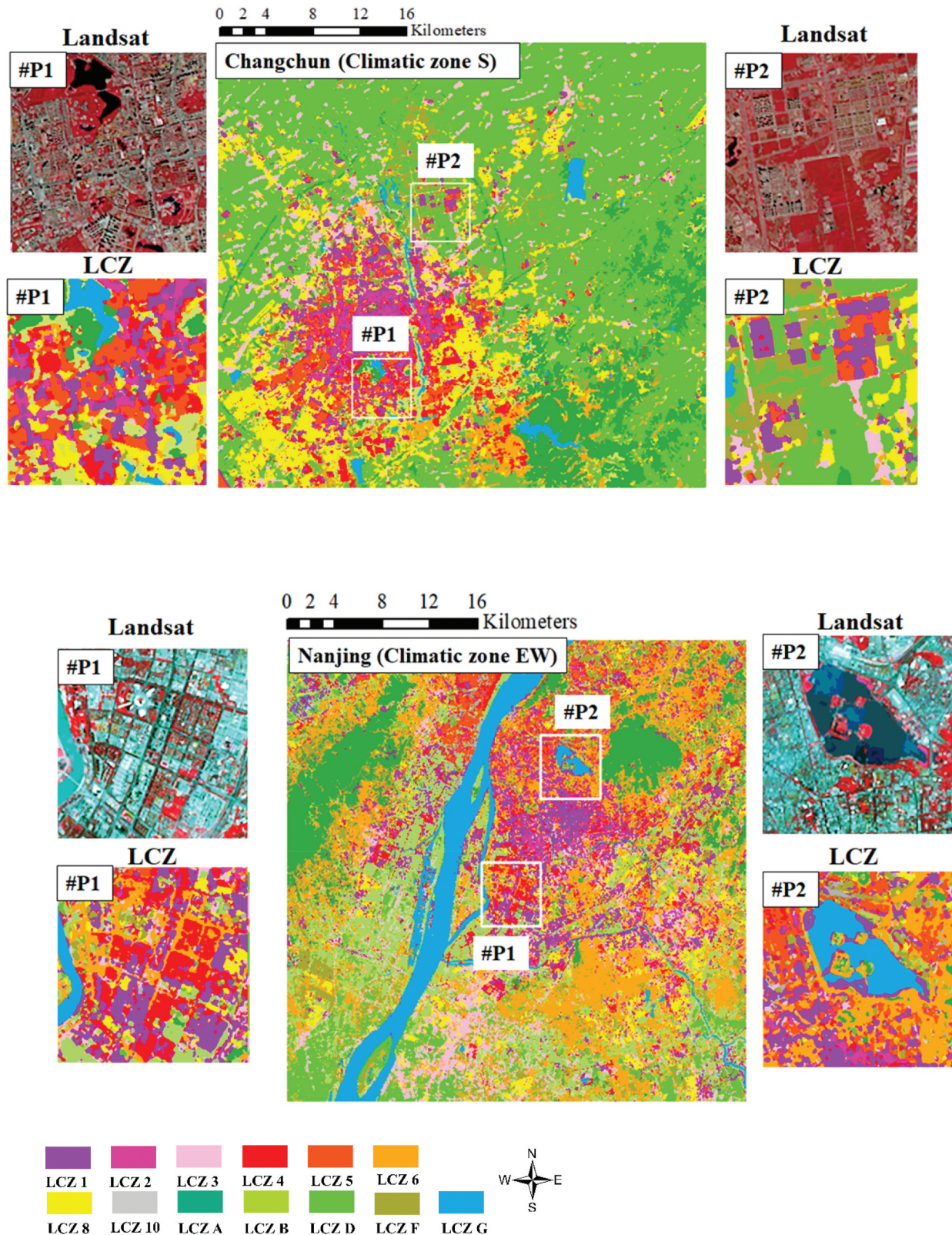


Figure 9. (Continued).

and change detection (Wang, Lu, and Qin 2020). The MLP has the advantage of data-driven and automatic learning, which is conducive to specific classification tasks. However, the MLP usually learns the non-linear spectral feature space at the pixel level, without considering the spatial context implicit in the images (Zhang et al. 2018).

Differing from all the aforementioned classifiers, the LCZ-CNN model is a data-driven feature learning and classification method (Levine et al. 2016). It also does not need us to set handcrafted image descriptors (e.g. GLCM, Gabor, Scale-Invariant Feature Transform (SIFT)) in advance. The network parameters can be adjusted during the training process to

**Table 1.** Comparison of the computation times of WUDAPT, BoVW, MLP, and LCZ-CNN.

No. of samples	Computation time (s)					
	WUDAPT		BoVW		MLP	LCZ-CNN
	Spec	Spec+GLCM	Spec	Spec+GLCM		
10000	5.33	35.26	35.65	1286.81	17.17	22.18
50000	26.05	163.79	120.91	6725.42	97.14	107.03

learn the spectral and spatial features applicable to a specific interpretation task. Therefore, the convolutional layers of the proposed LCZ-CNN model can extract the spatial context information from Landsat images, and can learn more robust and adaptive features than the handcrafted descriptors (Zhang et al. 2018). Our experimental results demonstrate the superiority of the proposed LCZ-CNN model in mining and utilizing the spatial information contained in Landsat images.

Table 1 compares the computation time of the LCZ-CNN model with that of WUDAPT, BoVW, and MLP, for predicting 10,000 and 50,000 samples, respectively. These experiments were implemented using MATLAB 2016b with an Intel(R) Core(TM) i5-7500 CPU and 24.0 GB memory. The computation time of BoVW (Spec + GLCM) is the longest among all the methods. Its time cost mainly arises from the process of assigning local features to the corresponding visual words, the calculation of the frequency histogram of visual words used for the scene classification, and the GLCM feature extraction. It can be said that the efficiency of the LCZ-CNN model is significantly higher than that of BoVW (Spec), BoVW (Spec + GLCM), and WUDAPT (Spec + GLCM), when considering both the computation time and accuracy. The time complexity of a CNN is mainly related to the number of convolutional layers, the parameters of the convolutional kernels, and the channels of the feature maps generated by each convolutional layer (Wu et al. 2018). In this research, the proposed LCZ-CNN model was made up of only three convolutional stages and a GAP layer as a replacement for the fully connected layers in the conventional CNNs. Therefore, compared with the commonly used classical networks, e.g. AlexNet, Google Net, and VGG-16, the complexity of LCZ-CNN is greatly reduced. In conclusion, the LCZ-CNN model is a lightweight CNN that can achieve both a satisfactory classification accuracy and an acceptable computational efficiency.

Table 2 records the Floating Point Operations (FLOPs), which are proportional to the model complexity, and the average OA ( $OA_{avg}$ ) of the classification results from all 32 cities. Several recent commonly used networks, i.e. LeNet-5 (Lecun et al. 1998), DenseNet-121 (Huang et al. 2017), and ResNet-50 (He et al. 2016), are compared with LCZ-CNN in Table 2. LeNet-5 requires only one input channel, while DenseNet-121 and ResNet-50 require three input channels. In order to meet the requirements of these networks with regard to input channels, the 32-channel images were compressed into one or three channels through the use of the Principal Component Analysis (PCA) algorithm, according to the operation in Xiao et al. (2017), the purpose of which was to compare the performance of each network fairly. We adopted two strategies to exploit the CNN models: 1) fine-tuning the pre-trained network; and 2) fully training the network from scratch. It can be clearly seen from Table 2 that the complexity of LCZ-CNN is far less than that of DenseNet and ResNet, and the classification accuracy of the LCZ-CNN model is much higher than that of all three commonly used networks, which proves the superiority of LCZ-CNN. Several other studies (Chen et al. 2016; Ma et al. 2018; Song et al. 2019) have also concluded that directly applying these original classical CNNs to remote sensing tasks might not obtain satisfactory results.

## 5.2. Function and transferability of the convolutional stages

The proposed LCZ-CNN model is composed of three typical convolutional stages named Stage-1, Stage-2, and Stage-3, respectively (Figure 3(c)). A typical convolutional stage of a CNN includes a convolution operation, an elementwise non-linear function (e.g. ReLU non-linearity), and a pooling operation (Figure 3(b)). To explore the function and transferability of

**Table 2.** Comparison of the FLOPs and average OA for the 32 cities with LeNet-5, DenseNet-121, ResNet-50, and LCZ-CNN.

	<i>LeNet</i>	<i>LeNet<sub>pre</sub></i>	<i>DenseNet</i>	<i>DenseNet<sub>pre</sub></i>	<i>ResNet</i>	<i>ResNet<sub>pre</sub></i>	<i>LCZ_CNN</i>
FLOPs	$4.23 \times 10^5$	$4.23 \times 10^5$	$2.88 \times 10^9$	$2.88 \times 10^9$	$4.12 \times 10^9$	$4.12 \times 10^9$	$6.66 \times 10^6$
$OA_{avg}$	0.43	0.44	0.44	0.46	0.43	0.45	0.80

*LeNet*: LeNet-5 model trained from scratch with LCZ training samples.

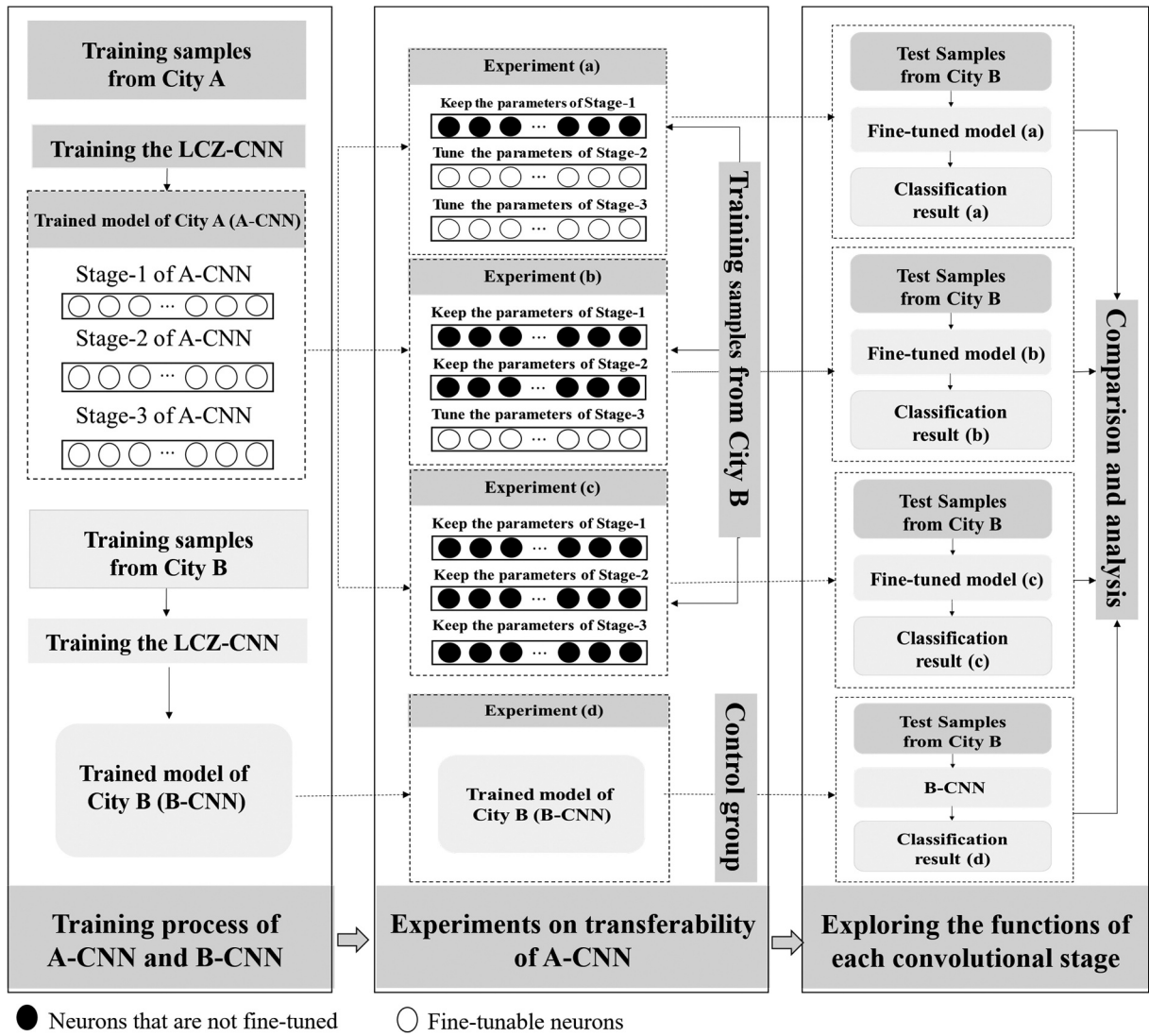
*LeNet<sub>pre</sub>*: Fine tuning the pre-trained LeNet-5 model of the MNIST dataset with LCZ training samples.

*DenseNet*: DenseNet-121 model trained from scratch with LCZ training samples.

*DenseNet<sub>pre</sub>*: Fine tuning the pre-trained DenseNet-121 model of the ImageNet dataset with LCZ training samples.

*ResNet*: ResNet-50 model trained from scratch with LCZ training samples.

*ResNet<sub>pre</sub>*: Fine tuning the pre-trained ResNet-50 model of the ImageNet dataset with LCZ training samples.



**Figure 10.** Experiments to explore the function and transferability of each convolutional stage in LCZ-CNN.

each convolutional stage in the LCZ-CNN model, we designed a number of experiments (Figure 10), as follows. Specifically, the city used to construct the original network was called City A, and the LCZ-CNN model trained by the samples from City A was denoted as A-CNN. Furthermore, the A-CNN model was fine-tuned by the samples from City B, in order to test the function and transferability of the different stages of the LCZ-CNN model. We therefore set a series of experiments:

**Experiment (a):** The parameters of Stage-1 of A-CNN were fixed, but the parameters of Stage-2 and Stage-3 were tuned according to the training samples from City B, and the fine-tuned model was applied to classify the images of city B. The purpose of this experiment was to explore the function and transferability of Stage-1.

**Experiment (b):** The parameters of Stage-1 and Stage-2 of A-CNN were fixed, and the parameters of Stage-3 were tuned with the training samples from City B. The fine-tuned model were then used for

City B. This experiment was aimed at investigating the transferability of the features extracted by Stage-2.

**Experiment (c):** The parameters of all three convolutional stages (Stage-1, Stage-2, and Stage-3) of A-CNN were fixed. This is equivalent to using A-CNN as a feature extractor to classify the test samples from City B, without consideration of the samples from City B. In this way, the transferability and function of Stage-3 could be analyzed.

**Experiment (d):** A new LCZ-CNN model, i.e. B-CNN, was trained from scratch using training samples from City B. This experiment was conducted as a comparison.

The purpose of the above experiments was to analyze the transferability of the parameters in the various convolutional stages for LCZ mapping. Specifically, we chose three city groups for transfer learning: Guangzhou to Beijing, Hohhot to Hefei, and Shenyang to Fuzhou (the former is City A, and the latter is City B). In the experiments, City A and City B were chosen from different climatic zones. Table 3



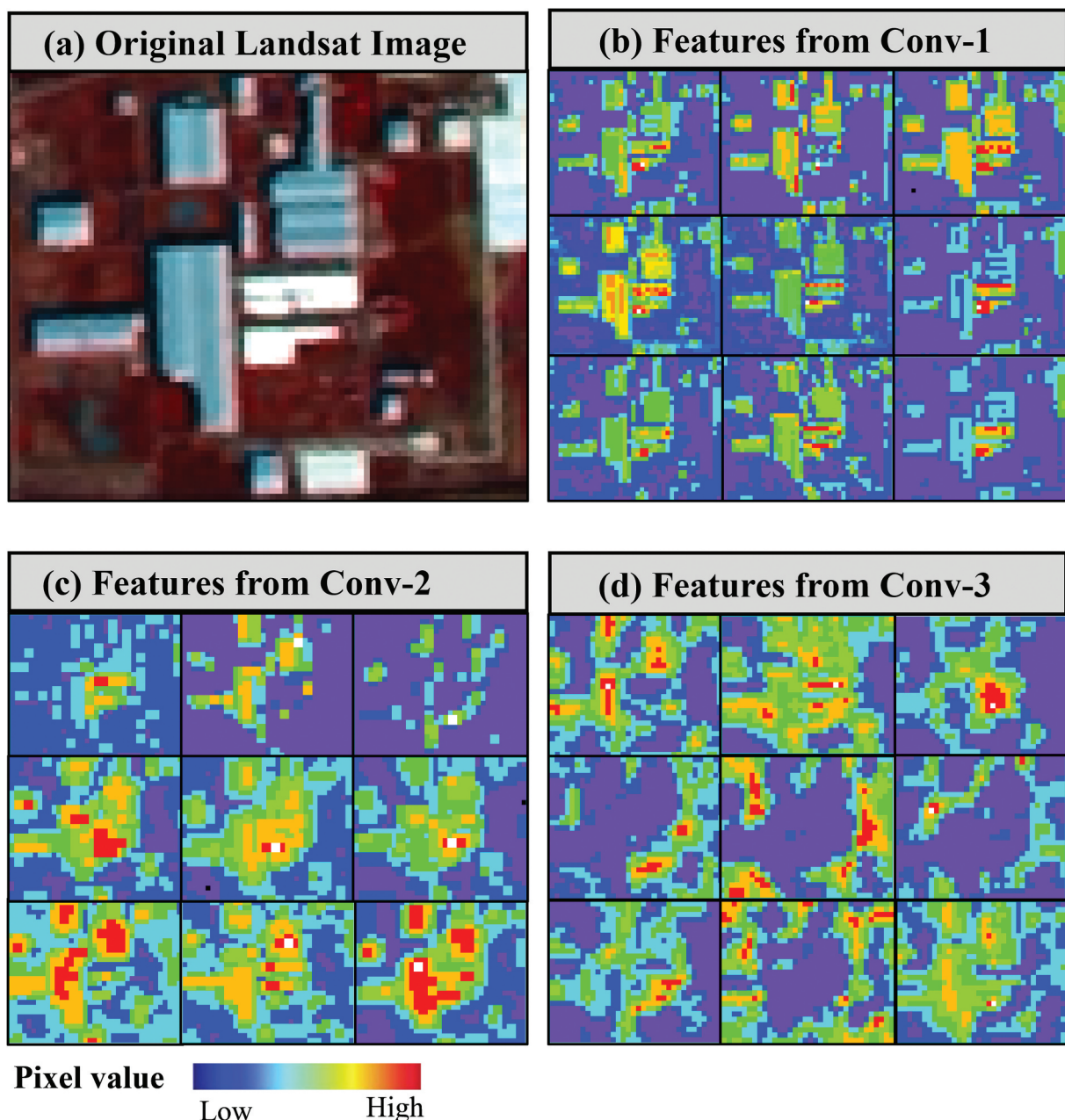
**Table 3.** Overall accuracies for the transfer learning experiments.

City A to City B	Guangzhou to Beijing	Hohhot to Hefei	Shenyang to Fuzhou
Experiment (a)	0.64	0.66	0.65
Experiment (b)	0.54	0.49	0.44
Experiment (c)	0.24	0.30	0.34
Experiment (d)	0.86	0.82	0.85

compares the results obtained by the three CNN models, i.e. the cases of (a), (b), and (c), with the result of the self-trained network (the case of (d)).

In Experiment (a), although the classification performances are inferior to those of Experiment (d), the OAs are greater than 60%. This implies that the first

convolutional stage can extract low-level visual features (e.g. spectra, texture), which are generally invariant between different cities, so that the parameters of Stage-1 demonstrate transferability, to some degree (Alshehhi et al. 2017). In Experiment (b), the LCZ mapping accuracy is significantly decreased when compared to Experiment (a), with OA values of <60%. However, the accuracy of Experiment (b) is superior to that of Experiment (c). The results of Experiment (b) and Experiment (c) indicate that the parameters of Stage-2 and Stage-3 possess poor transferability. The objects, as well as their spectral-spatial characteristics, present considerable differences between the different cities, and although they belong to the same LCZ category, there are notable



**Figure 11.** The original input image (LCZ 8, large low-rise) (a) and the features extracted from the three convolutional layers of LCZ-CNN (b)–(d). (b) The first convolutional layer. (c) The second convolutional layer. (d) The third convolutional layer. Nine feature maps arranged in a  $3 \times 3$  layout are selected for each convolutional layer.

distinctions in the image acquisition conditions (e.g. view angle, weather, atmosphere) and urban landscapes among the various cities (Yang and Lo 2000; Vogelmann et al. 2001). Consequently, the features extracted from Stage-2 and Stage-3 are probably more closely related to the specific image characteristics of the LCZs from the target area (Traviglia and Torsello 2017), but are not suitable for transfer to other cities. The function of Stage-2 and Stage-3 is to further extract and highlight the task-oriented image features on the basis of the low-level visual features extracted by Stage-1, which actually limits the generalization ability of the extracted features obtained in Stage-2 and Stage-3.

The representative features extracted from all three convolutional layers of LCZ-CNN are displayed in Figure 11. The three convolutional layers in LCZ-CNN are denoted as Conv-1, Conv-2, and Conv-3, respectively. Figure 11(b) highlights the shape and edges of the buildings in Figure 11(a), and demonstrates that the Conv-1 features retain the original contours of the ground objects. Figure 11(c) retains the fuzzy edges and geographical spatial relationships of the buildings in Figure 11(a,d) indicates the semantics of the buildings (or no buildings) in the urban scenes. The features extracted by Conv-3 are more abstract, and can be considered to be more likely to express higher-level semantic information (Liu et al. 2017).

Figure 12 presents the intra-class similarity and the inter-class separability in the three stages. The intra-class similarity was measured with Pearson correlation coefficients (Hauke and Kossowski 2011), and the inter-class separability between the different classes was calculated with the Jeffries-Matusita distance (Dabboor et al. 2014). Some LCZ categories (e.g. LCZ 1/LCZ 2 and LCZ 4/LCZ 5) that are difficult to distinguish were chosen in the analysis. With the increase of the number of convolutional stages, the receptive field becomes larger and the extracted features become more global and abstract (Figure 11).

Meanwhile, the inter-class separability exhibit a trend of growth (Figure 12(b)). It can be seen that the intra-class similarity is smaller in Stage-1 compared to the original images, but is gradually increased, and in the final stage, the intra-class similarity is larger than in the original images (Figure 12(a)). In summary, the above results illustrate that the features extracted by Stage-1 and Stage-2 are not as robust and discriminative as those of Stage-3. Therefore, it is necessary to set up Stage-3 to further convolute and integrate the features extracted in the first two convolutional stages.

### 5.3. The contribution of seasonal information

Some studies (Zhao et al. 2016; Zhu and Liu 2014) have pointed out that seasonal information can optimize the mapping results of thematic land-cover classification tasks based on Landsat images. However, the LCZ scheme is a newly developed concept and has intrinsic differences with the traditional thematic mapping tasks such as land-cover mapping, land-use mapping, and object detection (Wang, Lu, and Qin 2020). To investigate the contribution of seasonal information to the proposed LCZ-CNN model, we compared the OA of the LCZ classification obtained using multi-seasonal and single-seasonal inputs, respectively, for five representative cities (Shanghai, Taiyuan, Tianjin, Yinchuan and Lhasa) chosen across different climatic zones (Figure 13). The results show that, with the addition of multi-seasonal images, the OAs for the five representative cities are improved by 2.6%, 4.0%, 3.9%, 6.8%, and 10.9%, respectively, compared with the classification results based on only spring images. Specifically, taking Shanghai as an example, the OAs of the results for spring, summer, autumn, and winter are 70.9%, 68.9%, 53.4%, and 66.2%, respectively, but the accuracy obtained by considering all four seasonal images is increased significantly to 73.6%. The reason for this is that multi-temporal information can introduce the phenological characteristics of plants (Peng et al. 2009), and

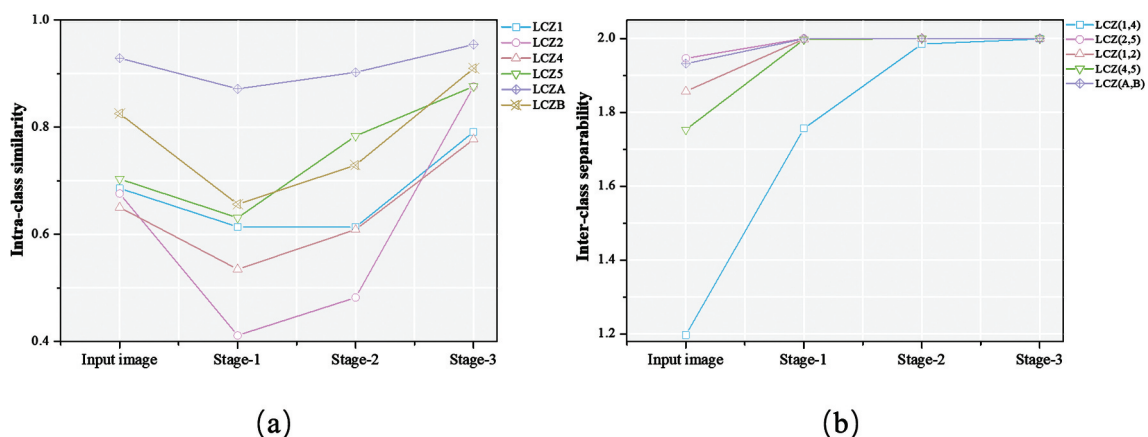
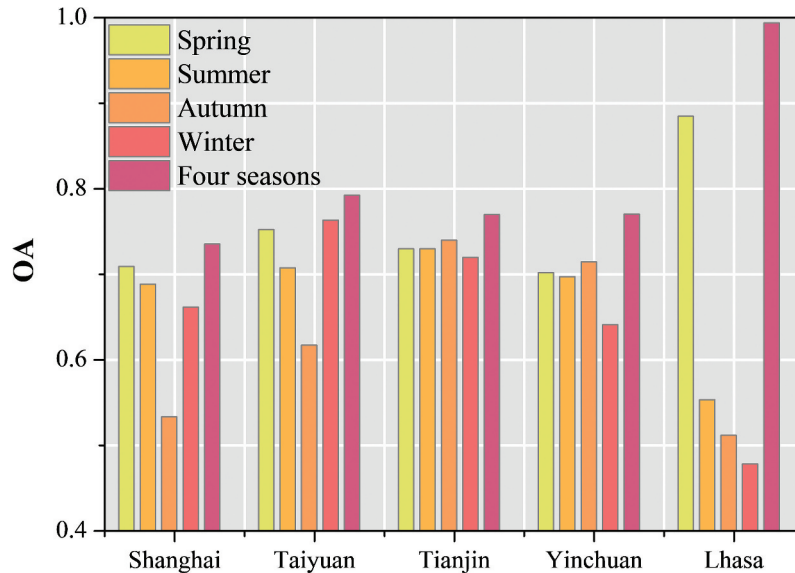


Figure 12. Intra-class similarity and inter-class separability of the extracted features.



**Figure 13.** Comparison of the overall accuracies obtained by LCZ-CNN based on multi-seasonal information and single-seasonal input.

a lot of LCZ classes, e.g. LCZ 4 (open high-rise), LCZ 5 (open mid-rise), LCZ 6 (open low-rise), LCZ A (dense trees), and LCZ B (scattered trees), refer to vegetation scenarios. Figure 13 also reveals that the contribution of each season is different in the different cities and climatic zones, and hence it is difficult to clearly identify which season is optimal in LCZ classification. To sum up, our experimental results show that it is appropriate to make full use of the multi-seasonal images and capture the temporal characteristics for LCZ classification.

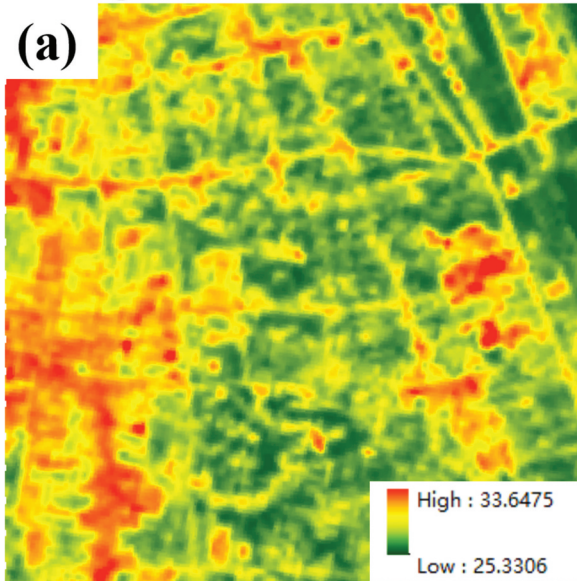
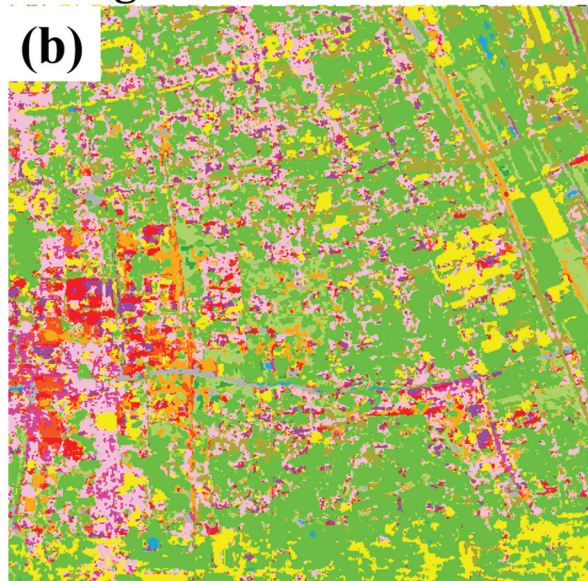
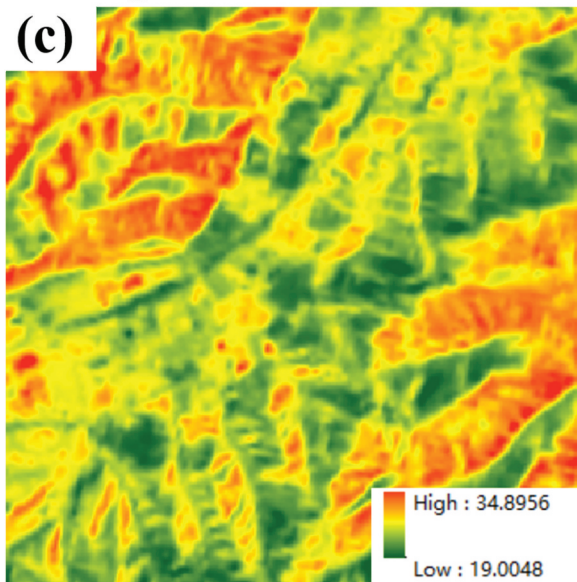
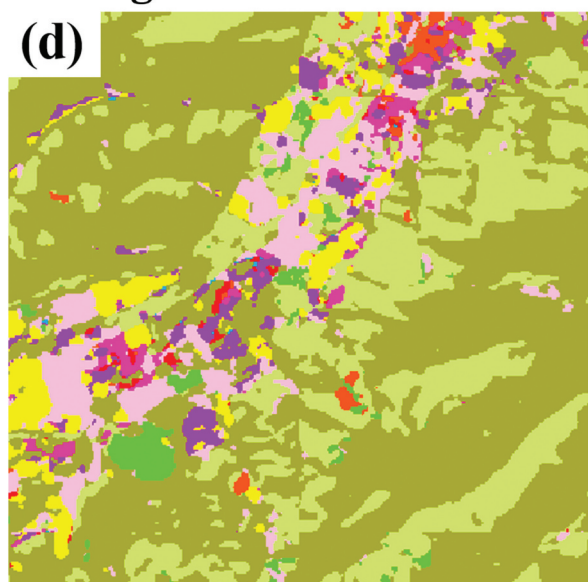
#### 5.4. Thermal characteristics of the LCZs in China's 32 major cities

The LCZ scheme is a standard framework used to describe urban structures, which can reflect the local temperature and microclimate of a city (Stewart and Oke 2012). For instance, Figure 14 shows the LCZ mapping results and the corresponding LST in the two cities of Shanghai and Xining, where it is apparent that the spatial morphology of the LCZs is closely correlated to that of the LST.

According to formula (4),  $\Delta LST_i$  can be calculated to represent the departure from the mean LST for the  $i_{th}$  LCZ class. The range of  $\Delta LST_i$  is divided into four classes at an interval of  $1^\circ\text{C}$ . The values of  $\Delta LST_i > 0$  refer to a heat island, and  $\Delta LST_i < 0$  represents a cold island. Figure 15 shows that the built types generally show a significant heat island effect ( $\Delta LST_i \geq 2^\circ\text{C}$ ), while the cold island effect in most cities is concentrated in LCZ A (dense trees), B (scattered trees), D (low plants), and G (water).

Figure 16 presents box plots showing the departure from the mean LST for all the LCZ classes in China's 32 major cities, which indicate significant thermal differences among the various LCZ classes. From the

figure, it can be found that LCZ A (dense trees) and LCZ G (water) have the lowest LST, while LCZ 2 (compact mid-rise), LCZ 3 (compact low-rise), and LCZ 8 (large low-rise) are among the highest. LCZ 2 (compact mid-rise) and LCZ 3 (compact low-rise) are usually located in bustling commercial or residential areas, while the majority of LCZ 8 (large low-rise) corresponds to industrial areas. The LST of most built types (LCZ 1–10) and LCZ F (bare soil or sand) is higher than the mean LST of all the LCZ classes, and the LST of the natural types (LCZ A, B, D, and G) is lower than the mean LST. This can be attributed to the large amount of heat generated from the consumption and re-radiation of solar radiation from urban structures and anthropogenic heat sources (Rizwan, Dennis, and Liu 2008). In particular, it is interesting to see that the compact built types (LCZ 1–3) generally have a higher LST than the open types (LCZ 4–6), since the former are mainly located in commercial and residential areas with high population density and frequent human activities. Moreover, the area proportion of the permeable layer, i.e. vegetation, water, and soil, is low in these compact built types, and the effect of evapotranspiration and shading by plants can significantly reduce the amount of heat that is re-radiated by urban constructions. Furthermore, compact buildings and narrow streets can form canyon structures, which increase the absorption of solar radiation, and lead to the trapping of long-wave radiation from the ground (Yuan and Ng 2012). The walls of buildings can also reduce wind speed and prevent air flow from cooling the street canyons, contributing to further heating of the land surface (Oke 1988). It is also found that the LST of the high-rise built types (LCZ 1, 4) is lower than that of the low- and mid-rise built types (LCZ 2, 3, 5, and 6). A reasonable explanation is

**Shanghai-LST****Shanghai-LCZ****Xining-LST****Xining-LCZ**

**Figure 14.** Selected regions in Shanghai and Xining, displaying the close relationship and similarity between LST ( $^{\circ}\text{C}$ ) and the LCZ mapping results.

that high-rise buildings can generate larger shaded areas than low- and mid-rise buildings, leading to lower LST in the neighborhood (Perini and Magliocco 2014). Another possible reason is that high-rise built types usually have a lower building density, which actually provides more open or green space, and hence decreases the LST (Cai and Xu 2017). In general, our results indicate that the spatial configuration of LCZs can significantly affect the magnitude of the LST.

Figure 16 is further decomposed into a series of sub-figures (Figure 17) in terms of the climatic zones. From Figure 17(a), (b), and (c), it can be observed that the thermal characteristics of climatic zones S, EW, and W are similar to the overall pattern in Figure 16, indicating that the built types generally have a higher LST than most of the natural types. However, the LST pattern of climatic zones A and TS is clearly different from the other three (Figure 17(d)), where LCZ F (bare soil or sand)

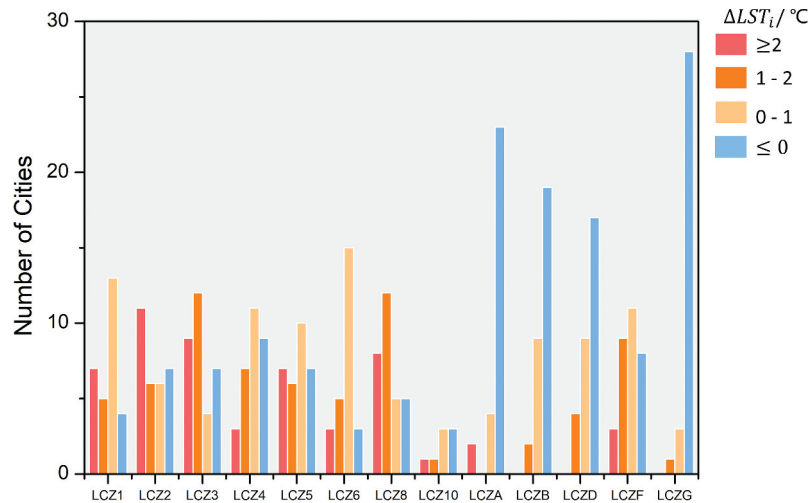


Figure 15.  $\Delta LST_i$  frequency distribution for the 32 cities.

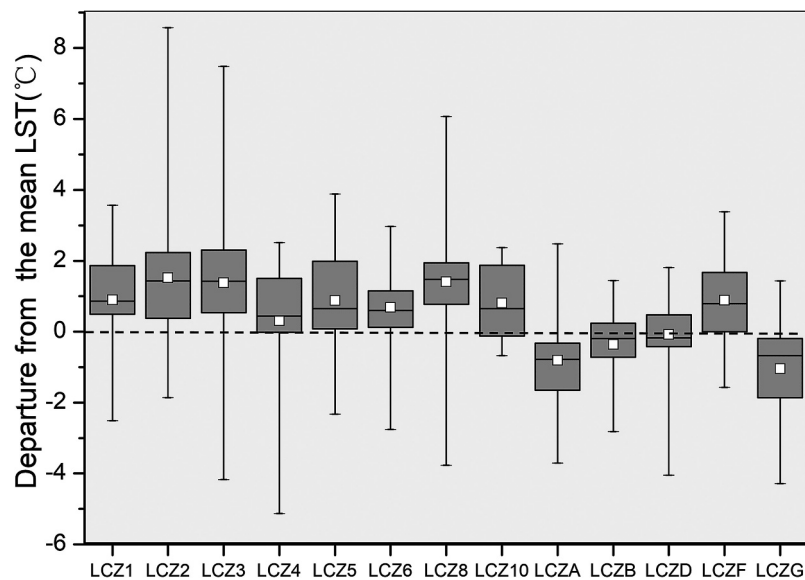


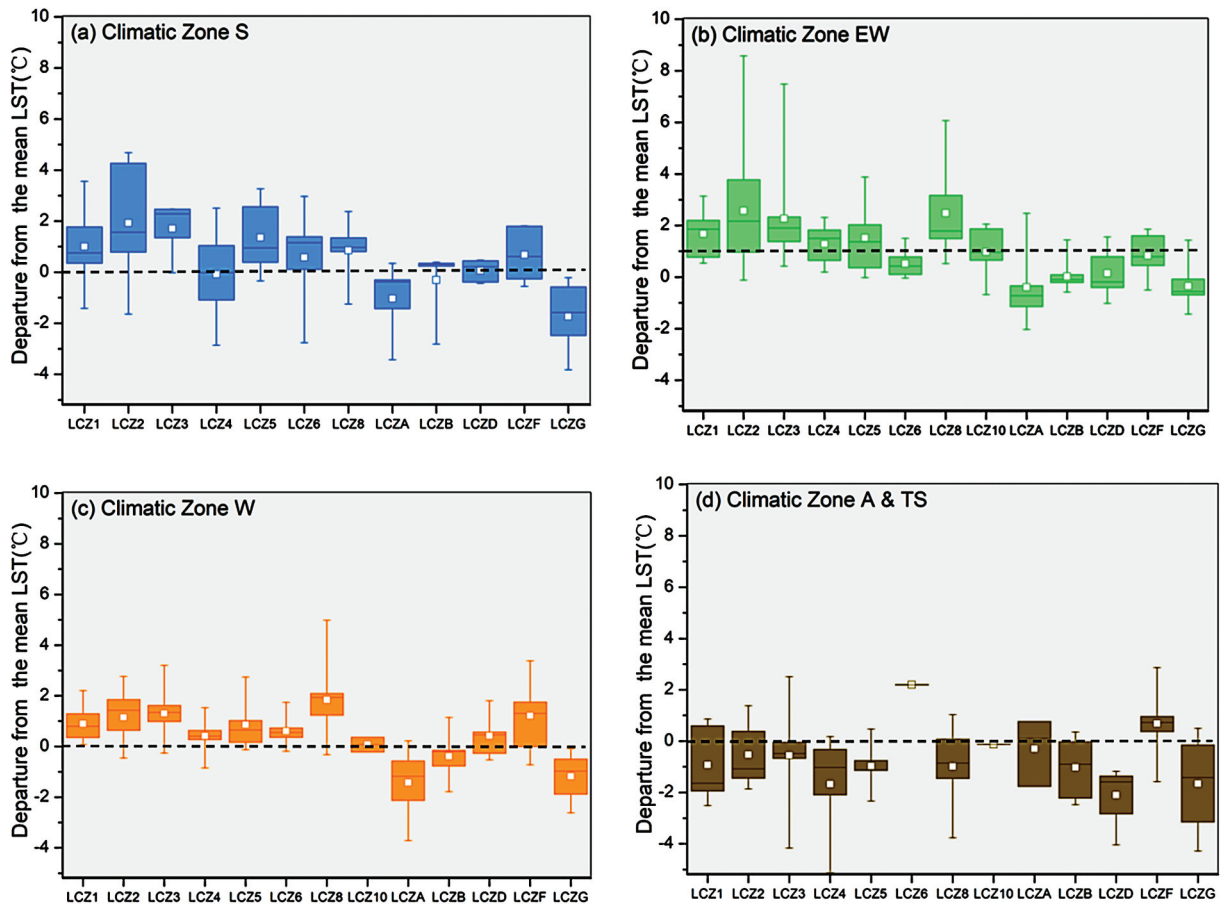
Figure 16. Departure from the mean LST for all the LCZ classes in China's 32 major cities.

generally has a higher LST than the mean LST for all the LCZ classes, while most built types have a lower LST than the average LST. As shown in Figure 8, the area proportion of LCZ F (bare soil or sand) is high in the cities of climatic zones A and TS. This is due to the fact that, in these climatic zones, desert occupies a large amount of space around the built-up areas of these cities (see Figure 9). Owing to the lower heat capacity of the desert, it warms up faster than the land cover in the built-up areas. The warm air heated over the desert is transported over the built-up areas through local circulation, forming an inversion layer of upper heat and lower cold, which keeps the lower cold air stable, thus forming a relatively cool and humid microclimate in the urban areas (Hu, Su, and Zhang 1988). Therefore, the LST of the built-up LCZ classes in the A and TS climatic zones is significantly lower than that in the other zones. Similar phenomena have emerged in studies of Phoenix in the Southwestern U.S. (Wang et al. 2018) and Dubai in the Middle East

(Nassar, Blackburn, and Whyatt 2016), both of which are typical cities surrounded by sand and soil, with an arid climate. However, differing from our results, there are still some LCZ classes in Phoenix with a higher LST than LCZ F (soil or sand), such as LCZ 8 (large low-rise). Our results for Western China are more similar to those in Dubai (Nassar, Blackburn, and Whyatt 2016), where LCZ F maintains a much higher LST in the daytime than all the other LCZ classes, which is perhaps due to the fact that both study areas are located in Asia.

## 6. Conclusion

This study was aimed at generating more accurate LCZ maps and investigating the relationship between LCZs and the thermal environment in China's 32 major cities. To this aim, in this paper, we have proposed a novel CNN framework (namely, LCZ-CNN) designed to cope with the issues of LCZ classification



**Figure 17.** Departure from the mean LST for all the LCZ classes of the different climatic zones. (a) Climatic zone S. (b) Climatic zone EW. (c) Climatic Zone W. (d) Climatic zone A&TS.

using Landsat imagery. The LCZ-CNN model adequately considers the multi-spectral and multi-seasonal information contained in Landsat images. To evaluate its performance, the proposed method was tested over China's 32 major cities distributed in different climatic zones, and the traditional classification strategies of WUDAPT, BoVW, and MLP were also applied for comparison. The proposed LCZ-CNN model produced the most accurate mapping results and a satisfactory computational efficiency in all the study areas. A systematic analysis of the proposed LCZ-CNN model was conducted, including the transferability of the network and the effectiveness of multi-seasonal information. It was found that the first convolutional stage, corresponding to low-level features, shows better transferability than the second and third convolutional stages, which correspond to high-level and task-oriented features. It was also found that multi-seasonal information can significantly improve the accuracy of LCZ classification.

The relationship between LCZs and their thermal properties was investigated in detail over China's 32 major cities. To the best of our knowledge, this is the first time that this relationship has been analyzed on the basis of a large number of

cities across multiple climatic zones. Overall, it was found that the compact built types have higher LSTs than the open classes, and the LSTs of the high-rise built types are lower than those of the low- and mid-rise built types. Our experimental results confirmed the close relationship between the LCZ classes and the magnitude of the LST. For most cities in climatic zones S (snow climate with dry winter), W (warm temperature climate with dry winter), and EW (equatorial climate and warm and fully humid temperate climate), most built types have a significantly higher LST than their neighborhoods, and hence show the heat island effect. In contrast, the LCZs in climatic zones A (climate of arid steppe and desert) and TS (tundra climate and snow climate with cool summer and cold winter) demonstrate distinct thermal characteristics, i.e. LCZ F (bare soil or sand) has the highest LST, while most built types have a lower LST than the average. This phenomenon can be attributed to the unique urban landscape and climate conditions in the arid cities of China.

In the future, we will work toward constructing more powerful deep networks which have superior

transfer capabilities and can obtain higher accuracies. The interpretation of LST from the LCZ viewpoint should also be further investigated by taking into account more cities across the globe. In addition, although Landsat images have been widely used for LCZ mapping in the existing literature, thanks to their free and convenient access and adequate data archives, high-resolution remotely sensed imagery should be considered in the future, since such imagery can provide more spatial and structural information, and has the potential to be more effective in delineating the composition and configuration of the various LCZ classes.

### Data availability statement

The data that support the findings of this study are available from the corresponding author [X. Huang] at [http://irsip.whu.edu.cn/resources/resources\\_en\\_v2.php](http://irsip.whu.edu.cn/resources/resources_en_v2.php), upon reasonable request.

### Funding

This work is supported by the National Natural Science Foundation of China [Grant Nos 41971295 and 42071311].

### Notes on contributors

**Xin Huang** received the PhD degree in photogrammetry and remote sensing in 2009 from Wuhan University, Wuhan, China. He is currently a Luojia Distinguished Professor of Wuhan University, Wuhan, China, where he teaches remote sensing, photogrammetry, image interpretation, etc. He is the Founder and Director of the Institute of Remote Sensing Information Processing (IRSIP), School of Remote Sensing and Information Engineering, Wuhan University. He has published more than 150 peer-reviewed articles in the international journals. His research interests include remote sensing image processing methods and applications. He was the recipient of the Boeing Award for the Best Paper in Image Analysis and Interpretation from the American Society for Photogrammetry and Remote Sensing (ASPRS) in 2010, the second place recipient for the John I. Davidson President's Award from ASPRS in 2018, and the National Excellent Doctoral Dissertation Award of China in 2012. He was an Associate Editor of the *Photogrammetric Engineering and Remote Sensing* (2016–2019), and now serves as an Associate Editor of the *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS* (since 2014), and an Associate Editor of the *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING* (since 2018). He is also an editorial board member of some international journals in the field of remote sensing and the geo-spatial information science.

**Anling Liu** is currently pursuing the MS degree at the School of Remote Sensing and Information Engineering, Wuhan University. Her research interests include land cover/ land use classification and deep learning.

**Jiayi Li** received the PhD degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2016. She is currently an Assistant Professor with the School

of Remote Sensing and Information Engineering, Wuhan University. She has authored over 30 peer-reviewed articles in the international journals. Her research interests include hyperspectral imagery, sparse representation, computation vision and pattern recognition, and remote sensing images.

### ORCID

Xin Huang  <https://orcid.org/0000-0002-5625-0338>

Anling Liu  <https://orcid.org/0000-0002-1072-063X>

Jiayi Li  <https://orcid.org/0000-0003-1691-9743>

### References

- Alexander, P., and G. Mills. 2014. "Local Climate Classification and Dublin's Urban Heat Island." *Atmosphere* 5 (4): 755–774. doi:10.3390/atmos5040755.
- Alshehhi, R., P. R. Marpu, W. L. Woon, and M. D. Mura. 2017. "Simultaneous Extraction of Roads and Buildings in Remote Sensing Imagery with Convolutional Neural Networks." *ISPRS Journal of Photogrammetry and Remote Sensing* 130: 139–149. doi:10.1016/j.isprsjprs.2017.05.002.
- Arun, K. S., V. K. Govindan, and S. D. M. Kumar. 2020. "Enhanced Bag of Visual Words Representations for Content Based Image Retrieval: A Comparative Study." *Artificial Intelligence Review* 53 (3): 1615–1653. doi:10.1007/s10462-019-09715-6.
- Aslam, M. A., M. N. Salik, F. Chughtai, N. Ali, S. H. Dar, and T. Khalil. 2019. "Image Classification Based on Mid-Level Feature Fusion." Paper presented at the 2019 15th International Conference on Emerging Technologies (ICET), December 2–3. doi:10.1109/ICET48972.2019.8994721.
- Bechtel, B., P. Alexander, J. Böhner, J. Ching, O. Conrad, J. Feddema, G. Mills, L. See, and I. Stewart. 2015. "Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities." *ISPRS International Journal of Geo-Information* 4: 199–219. doi:10.3390/ijgi4010199.
- Bechtel, B., M. Demuzere, G. Mills, W. Zhan, P. Sismanidis, C. Small, and J. Voogt. 2019. "SUHI Analysis Using Local Climate Zones—A Comparison of 50 Cities." *Urban Climate* 28: 100451. doi:10.1016/j.uclim.2019.01.005.
- Cai, H., and X. Xu. 2017. "Impacts of Built-Up Area Expansion in 2D and 3D on Regional Surface Temperature." *Sustainability* 9: 1862. doi:10.3390/su9101862.
- Cai, M., C. Ren, Y. Xu, K. K.-L. Lau, and R. Wang. 2018. "Investigating the Relationship between Local Climate Zone and Land Surface Temperature Using an Improved WUDAPT Methodology – A Case Study of Yangtze River Delta, China." *Urban Climate* 24: 485–502. doi:10.1016/j.uclim.2017.05.010.
- Chen, S., H. Wang, F. Xu, Y. Q. Jin, and R. Sensing. 2016. "Target Classification Using the Deep Convolutional Networks for SAR Images." *IEEE Transactions on Geoscience* 54 (8): 4806–4817.
- Ching, J., G. Mills, L. See, B. Bechtel, J. Feddema, A. Hanna, G. Milcinski, et al. 2015. "The Portal Component, Strategic Perspectives and Review of Tactical Plans for Full Implementation of WUDAPT." Paper presented at the International Conference on Urban Climate 9 - ICUC9, July 21. doi:10.1016/j.uclim.2017.07.007.
- Dabboor, M., S. Howell, M. Shokr, and J. Yackel. 2014. "The Jeffries–Matusita Distance for the Case of Complex

- Wishart Distribution as a Separability Criterion for Fully Polarimetric SAR Data.” *International Journal of Remote Sensing* 35 (19): 6859–6873. doi:10.1080/01431161.2014.960614.
- Danylo, O., L. See, B. Bechtel, D. Schepaschenko, and S. Fritz. 2016. “Contributing to WUDAPT: A Local Climate Zone Classification of Two Cities in Ukraine.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (5): 1841–1853. doi:10.1109/JSTARS.2016.2539977.
- Fisher, J. I., J. F. Mustard, and M. A. Vadeboncoeur. 2006. “Green Leaf Phenology at Landsat Resolution: Scaling from the Field to the Satellite.” *Remote Sensing of Environment* 100 (2): 265–279. doi:10.1016/j.rse.2005.10.022.
- Foody, G. M. 2002. “Status of Land Cover Classification Accuracy Assessment.” *Remote Sensing of Environment* 80 (1): 185–201. doi:10.1016/S0034-4257(01)00295-4.
- Gao, Y., J. Shi, J. Li, and R. Wang. 2021. “Remote Sensing Scene Classification Based on High-order Graph Convolutional Network.” *European Journal of Remote Sensing*. Advance online publication. doi:10.1080/22797254.2020.1868273.
- Geletič, J., and M. Lehnert. 2016. “GIS-based Delineation of Local Climate Zones: The Case of Medium-sized Central European Cities.” *Moravian Geographical Reports* 24: 2–12. doi:10.1515/mgr-2016-0012.
- Gong, P., S. Liang, E. J. Carlton, Q. Jiang, J. Wu, L. Wang, and J. V. Remais. 2012. “Urbanisation and Health in China.” *The Lancet* 379 (9818): 843–852. doi:10.1016/S0140-6736(11)61878-3.
- Hauke, J., and T. Kossowski. 2011. “Comparison of Values of Pearson’s and Spearman’s Correlation Coefficients on the Same Sets of Data.” *Quaestiones Geographicae* 30: 87–93. doi:10.2478/v10117-011-0021-1.
- He, K. M., X. Y. Zhang, S. Q. Ren, and J. Sun, and Ieee. 2016. “Deep Residual Learning for Image Recognition.” In *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. New York: IEEE. doi:10.1109/cvpr.2016.90.
- Hsiao, T.-Y., Y.-C. Chang, -H.-H. Chou, and C.-T. Chiu. 2019. “Filter-based Deep-compression with Global Average Pooling for Convolutional Networks.” *Journal of Systems Architecture* 95: 9–18. doi:10.1016/j.sysarc.2019.02.008.
- Hu, X. 2011. “Estimating Impervious Surfaces from Medium Spatial Resolution Imagery: A Comparison between Fuzzy Classification and LSMA.” *International Journal of Remote Sensing* 32: 5645–5663. doi:10.1080/01431161.2010.507258.
- Hu, Y., C. Su, and Y. Zhang. 1988. “Research on the Microclimate Characteristics and Cold Island Effect over a Reservoir in the Hexi Region.” *Advances in Atmospheric Sciences* 5 (1): 117–126. doi:10.1007/BF02657351.
- Huang, G., Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. 2017. “Densely Connected Convolutional Networks.” Paper presented at the 30th Ieee Conference on Computer Vision and Pattern Recognition. New York, July 21–26. doi:10.1109/cvpr.2017.243.
- Huang, X., X. Han, S. Ma, T. Lin, and J. Gong. 2019. “Monitoring Ecosystem Service Change in the City of Shenzhen by the Use of High-resolution Remotely Sensed Imagery and Deep Learning.” *Land Degradation & Development* 30 (12): 1490–1501. doi:10.1002/ldr.3337.
- Huang, X., S. Li, J. Li, X. Zhu, and J. Benediktsson. 2020. “A Multi-Spectral and Multi-Angle 3-D Convolutional Neural Network for the Classification of ZY-3 Satellite Images over Urban Areas.” *IEEE Transactions on Geoscience and Remote Sensing*. Advance online publication. doi:10.1109/TGRS.2020.3037211.
- Jamali, A. 2020. “Land Use Land Cover Mapping Using Advanced Machine Learning Classifiers: A Case Study of Shiraz City, Iran.” *Earth Science Informatics* 13 (4): 1015–1030. doi:10.1007/s12145-020-00475-4.
- Kemker, R., C. Salvaggio, and C. Kanan. 2018. “Algorithms for Semantic Segmentation of Multispectral Remote Sensing Imagery Using Deep Learning.” *ISPRS Journal of Photogrammetry and Remote Sensing* 145: 60–77. doi:10.1016/j.isprsjprs.2018.04.014.
- Krizhevsky, A., I. Sutskever, and G. E. Hinton. 2017. “Imagenet Classification with Deep Convolutional Neural Networks.” *Advances in Neural Information Processing Systems* 60 (6): 84–90. doi:10.1145/3065386.
- Lai, X., L. Zhou, Z. Fu, S. Naqvi, and J. Chambers. 2019. “Enhanced Pooling Method for Convolutional Neural Networks Based on Optimal Search Theory.” *IET Image Processing* 13 (12): 2152–2161. doi:10.1049/iet-ipr.2018.6322.
- Lau, K., C. Ren, Y. Shi, V. Zheng, S. H. L. Yim, and D. Lai. 2015. “Determining the Optimal Size of Local Climate Zones for Spatial Mapping in High-density Cities.” Paper presented at the Conference: The 9th International Conference on Urban Climate (ICUC9), Toulouse, France, July 20–24. [https://www.researchgate.net/profile/Kevin\\_Lau4/publication/280571730](https://www.researchgate.net/profile/Kevin_Lau4/publication/280571730).
- Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. “Gradient-Based Learning Applied to Document Recognition.” *Proceedings of the IEEE* 86: 2278–2324. doi:10.1109/5.726791.
- Lelovics, E., U. Janos, T. Gál, and C. Gal. 2014. “Design of an Urban Monitoring Network Based on Local Climate Zone Mapping and Temperature Pattern Modelling.” *Climate Research* 60: 51–62. doi:10.3354/cr01220.
- Levine, S., P. Pastor, A. Krizhevsky, J. Ibarz, and D. Quillen. 2016. “Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection.” *International Journal of Robotics Research* 37 (4–5): 421–436. doi:10.1007/978-3-319-50115-4\_16.
- Li, J., X. Huang, and J. Gong. 2019. “Deep Neural Network for Remote-sensing Image Interpretation: Status and Perspectives.” *National Science Review* 6 (6): 26–30. doi:CNKI:SUN:NASR.0.2019-06-011.
- Li, J., C. Song, L. Cao, F. Zhu, X. Meng, and J. Wu. 2011. “Impacts of Landscape Structure on Surface Urban Heat Islands: A Case Study of Shanghai, China.” *Remote Sensing of Environment* 115 (12): 3249–3263. doi:10.1016/j.rse.2011.07.008.
- Li, Y., N. Wang, J. Shi, X. Hou, and J. Liu. 2018. “Adaptive Batch Normalization for Practical Domain Adaptation.” *Pattern Recognition* 80: 109–117. doi:10.1016/j.patcog.2018.03.005.
- Liu, Y., S. Piramanayagam, S. T. Monteiro, and E. Saber. 2017. “Dense Semantic Labeling of Very-High-Resolution Aerial Imagery and LiDAR with Fully-Convolutional Neural Networks and Higher-Order CRFs.” Paper presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21–26. doi:10.1109/CVPRW.2017.200.
- Ma, X., A. Fu, J. Wang, H. Wang, and B. Yin. 2018. “Hyperspectral Image Classification Based on Deep Deconvolution Network with Skip Architecture.” *IEEE Transactions on Geoscience and Remote Sensing* 56 (8): 4781–4791. doi:10.1109/TGRS.2018.2837142.
- Meng, X., Y. Xiong, F. Shao, H. Shen, W. Sun, G. Yang, Q. Yuan, R. Fu, and H. Zhang. 2020. “A Large-Scale



- Benchmark Data Set for Evaluating Pansharpening Performance: Overview and Implementation.” *IEEE Geoscience and Remote Sensing Magazine*. Advance online publication. <https://www.researchgate.net/publication/340261023>
- Muhammad, U., W. Wang, A. Hadid, and S. Pervez. 2019. “Bag of Words KAZE (Bowk) with Two-step Classification for High-resolution Remote Sensing Images.” *Computer Vision, IET* 13 (4): 395–403. doi:10.1049/iet-cvi.2018.5069.
- Nassar, A. K., G. A. Blackburn, and J. D. Whyatt. 2016. “Dynamics and Controls of Urban Heat Sink and Island Phenomena in a Desert City: Development of a Local Climate Zone Scheme Using Remotely-sensed Inputs.” *International Journal of Applied Earth Observation and Geoinformation* 51: 76–90. doi:10.1016/j.jag.2016.05.004.
- Oke, T. R. 1988. “Street Design and Urban Canopy Layer Climate.” *Energy and Buildings* 11 (1): 103–113. doi:10.1016/0378-7788(88)90026-6.
- Otazu, X., M. González-Audicana, O. Fors, and J. Murga. 2005. “Introduction of Sensor Spectral Response into Image Fusion Methods. Application to Wavelet-Based Methods.” *IEEE Transactions on Geoscience and Remote Sensing* 43 (10): 2376–2385. doi:10.1109/TGRS.2005.856106.
- Peng, G., A. Gong, W. Cui, T. Min, and F. Chen. 2009. “Study on Methods Comparison of Typical Remote Sensing Classification Based on Multi-temporal Images: Study on Methods Comparison of Typical Remote Sensing Classification Based on Multi-temporal Images.” *Geo-information Science* 11: 225–230. doi:10.3724/SP.J.1047.2009.00225.
- Perini, K., and A. Magliocco. 2014. “Effects of Vegetation, Urban Density, Building Height, and Atmospheric Conditions on Local Temperatures and Thermal Comfort.” *Urban Forestry & Urban Greening* 13 (3): 495–506. doi:10.1016/j.ufug.2014.03.003.
- Rizwan, A. M., L. Y. C. Dennis, and C. Liu. 2008. “A Review on the Generation, Determination and Mitigation of Urban Heat Island.” *Journal of Environmental Sciences* 20 (1): 120–128. doi:10.1016/S1001-0742(08)60019-4.
- Roeser, S., R. Hillerbrand, P. Sandin, and M. Peterson. 2012. “Risk and Spatial Planning”. Chap.11. In *Handbook of Risk Theory*, edited by S. Roeser, 265–294. doi:10.1007/978-94-007-1433-5\_11.
- Simonyan, K., and A. Zisserman. 2014. “Very Deep Convolutional Networks for Large-Scale Image Recognition.” arXiv:1409.1556.
- Song, S., H. Yu, Z. Miao, Q. Zhang, Y. Lin, and S. Wang. 2019. “Domain Adaptation for Convolutional Neural Networks-Based Remote Sensing Scene Classification.” *IEEE Geoscience and Remote Sensing Letters* 16 (8): 1324–1328. doi:10.1109/LGRS.2019.2896411.
- Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. 2014. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting.” *Journal of Machine Learning Research* 15: 1929–1958.
- Stehman, S. V. 2013. “Estimating Area from an Accuracy Assessment Error Matrix.” *Remote Sensing of Environment* 132: 202–211. doi:10.1016/j.rse.2013.01.016.
- Stewart, I., T. Oke, and E. Krayenhoff. 2014. “Evaluation of the ‘Local Climate Zone’ Scheme Using Temperature Observations and Model Simulations.” *International Journal of Climatology* 34 (4): 1062–1080. doi:10.1002/joc.3746.
- Stewart, I. D., and T. Oke. 2012. “Local Climate Zones for Urban Temperature Studies.” *Bulletin of the American Meteorological Society* 93: 1879–1900. doi:10.1175/BAMS-D-11-00019.1.
- Streutker, D. R. 2003. “Satellite-measured Growth of the Urban Heat Island of Houston, Texas.” *Remote Sensing of Environment* 85 (3): 282–289. doi:10.1016/S0034-4257(03)00007-5.
- Subiyanto, S., A. Sukmono, N. Bashit, and F. J. Amarrohman. 2019. “The Use of a MLP Neural Network for Analysis and Aodeling of Land Use Changes with Variations Variable of Physical and Economic Social.” Paper presented at the IOP Conference Series: Earth and Environmental Science, December 13. doi:10.1088/1755-1315/389/1/012029.
- Szegedy, C., W. Liu, Y. Q. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. 2015. “Going Deeper with Convolutions.” Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, New York, July 12. doi:10.1109/cvpr.2015.7298594.
- Traviglia, A., and A. Torsello. 2017. “Landscape Pattern Detection in Archaeological Remote Sensing.” *Geosciences* 7: 128. doi:10.3390/geosciences7040128.
- Vogelmann, J. E., D. Helder, R. Morfitt, M. J. Choate, J. W. Merchant, and H. Bulley. 2001. “Effects of Landsat 5 Thematic Mapper and Landsat 7 Enhanced Thematic Mapper Plus Radiometric and Geometric Calibrations and Corrections on Landscape Characterization.” *Remote Sensing of Environment* 78 (1): 55–70. doi:10.1016/S0034-4257(01)00249-8.
- Wang, C., A. Middel, S. W. Myint, S. Kaplan, A. J. Brazel, and J. Lukasczyk. 2018. “Assessing Local Climate Zones in Arid Cities: The Case of Phoenix, Arizona and Las Vegas, Nevada.” *ISPRS Journal of Photogrammetry and Remote Sensing* 141: 59–71. doi:10.1016/j.isprsjprs.2018.04.009.
- Wang, M., Y. Q. Lu, and J. C. Qin. 2020. “A Dynamic MLP-based DDoS Attack Detection Method Using Feature Selection and Feedback.” *Computers & Security* 88 (1): 14. doi:10.1016/j.cose.2019.101645.
- Wicki, A., and E. Parlow. 2017. “Attribution of Local Climate Zones Using a Multitemporal Land Use/land Cover Classification Scheme.” *Journal of Applied Remote Sensing* 11: 026001. doi:10.1117/1.JRS.11.026001.
- Wu, X., R. He, Z. Sun, and T. Tan. 2018. “A Light CNN for Deep Face Representation with Noisy Labels.” *IEEE Transactions on Information Forensics and Security* 13 (11): 2884–2896. doi:10.1109/TIFS.2018.2833032.
- Xiao, Z., Y. Long, D. Li, C. Wei, G. Tang, and J. Liu. 2017. “High-Resolution Remote Sensing Image Retrieval Based on CNNs from a Dimensional Perspective.” *Remote Sensing* 9 (7): 725. doi:10.3390/rs9070725.
- Xu, Q., M. Zhang, Z. Gu, and G. Pan. 2019. “Overfitting Remedy by Sparsifying Regularization on Fully-connected Layers of CNNs.” *Neurocomputing* 328: 69–74. doi:10.1016/j.neucom.2018.03.080.
- Xu, Y., C. Ren, M. Cai, N. Y. Y. Edward, and T. Wu. 2017. “Classification of Local Climate Zones Using ASTER and Landsat Data for High-Density Cities.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10 (7): 3397–3405. doi:10.1109/JSTARS.2017.2683484.
- Yang, Q. Q., X. Huang, and J. Y. Li. 2017. “Assessing the Relationship between Surface Urban Heat Islands and Landscape Patterns across Climatic Zones in China.” *Scientific Reports* 7: 11. doi:10.1038/s41598-017-09628-w.
- Yang, X. J., and C. P. Lo. 2000. “Relative Radiometric Normalization Performance for Change Detection from Multi-date Satellite Images.” *Photogrammetric Engineering and Remote Sensing* 66 (8): 967–980. doi:10.1080/01431160050030628.
- Yang, Y., and S. Newsam. 2010. “Bag-of-visual-words and Spatial Extensions for Land-use Classification.” Paper presented at the Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic

- Information Systems, San Jose, California, November 2. doi:[10.1145/1869790.1869829](https://doi.org/10.1145/1869790.1869829).
- Yu, L., R. Zhang, S. Tian, L. Yang, and Y. Lv. 2018. "Deep Multi-Feature Learning for Water Body Extraction from Landsat Imagery." *Automatic Control and Computer Sciences* 52: 517–527. doi:[10.3103/S0146411618060123](https://doi.org/10.3103/S0146411618060123).
- Yuan, C., and E. Ng. 2012. "Building Porosity for Better Urban Ventilation in High-density Cities – A Computational Parametric Study." *Building and Environment* 50: 176–189. doi:[10.1016/j.buildenv.2011.10.023](https://doi.org/10.1016/j.buildenv.2011.10.023).
- Zhang, C., I. Sargent, X. Pan, H. Li, A. Gardiner, J. Hare, and P. Atkinson. 2018. "An Object-based Convolutional Neural Network (OCNN) for Urban Land Use Classification." *Remote Sensing of Environment* 216: 57–70. doi:[10.1016/j.rse.2018.06.034](https://doi.org/10.1016/j.rse.2018.06.034).
- Zhang, K., W. Zuo, Y. Chen, D. Meng, and L. Zhang. 2016. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising." *IEEE Transactions on Image Processing* 26 (7): 3142–3155. doi:[10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206).
- Zhang, W., X. He, and W. Lu. 2019. "Exploring Discriminative Representations for Image Emotion Recognition with CNNs." *IEEE Transactions on Multimedia* 22 (2): 515–523. doi:[10.1109/TMM.2019.2928998](https://doi.org/10.1109/TMM.2019.2928998).
- Zhao, Y., D. Feng, L. Yu, X. Wang, Y. Chen, Y. Bai, H. J. Hernández, et al. 2016. "Detailed Dynamic Land Cover Mapping of Chile: Accuracy Improvement by Integrating Multi-temporal Data." *Remote Sensing of Environment* 183: 170–185. doi:[10.1016/j.rse.2016.05.016](https://doi.org/10.1016/j.rse.2016.05.016).
- Zhu, X., and D. Liu. 2014. "Accurate Mapping of Forest Types Using Dense Seasonal Landsat Time-series." *ISPRS Journal of Photogrammetry and Remote Sensing* 96: 1–11. doi:[10.1016/j.isprsjprs.2014.06.012](https://doi.org/10.1016/j.isprsjprs.2014.06.012).

Appendix A

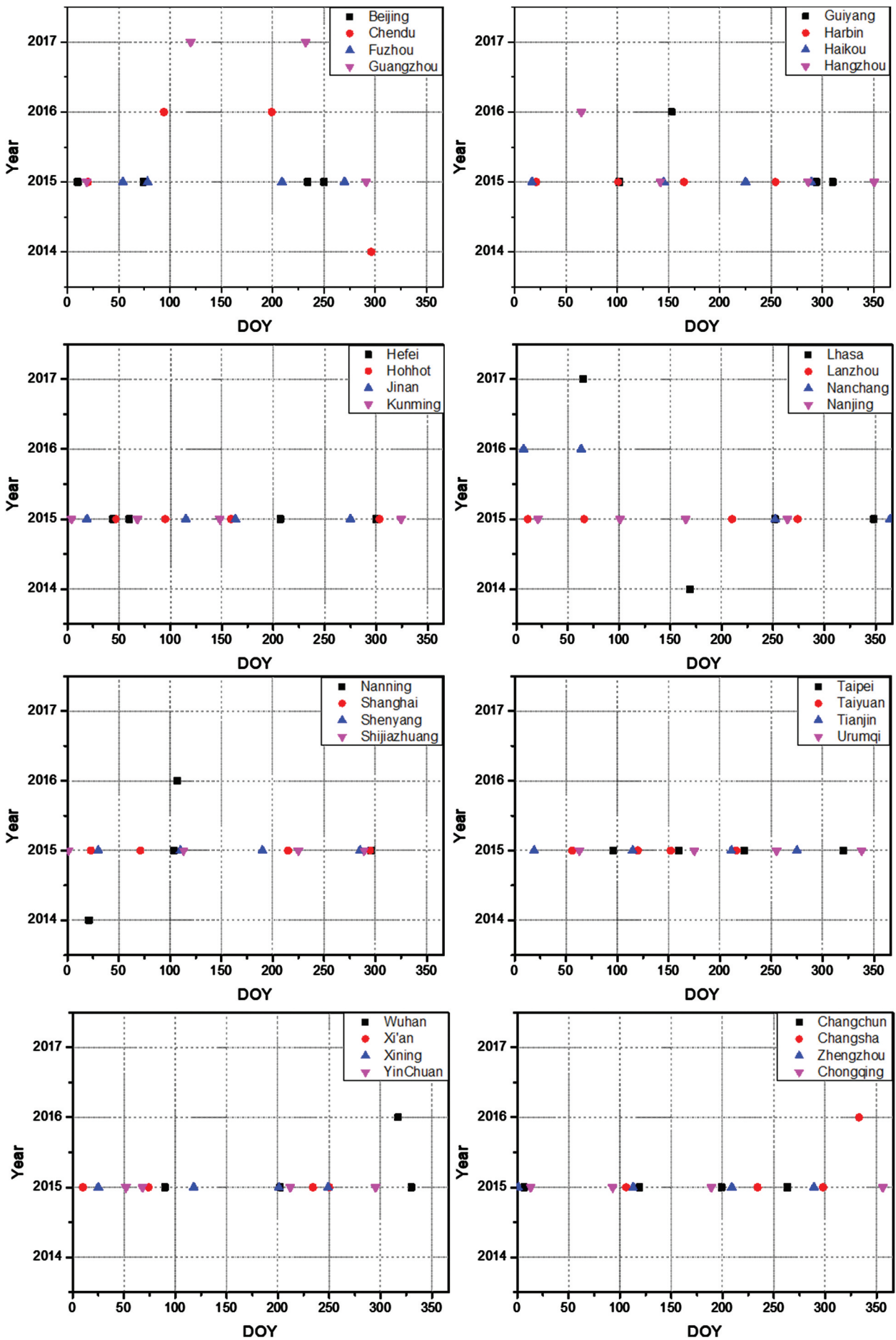


Figure A1. Landsat 8 images employed in this research. The vertical axis represents the acquisition year, and the horizontal axis represents the acquisition date in the form of "day of year" (DOY).

**Table A1.** Overall accuracies for different input window sizes of LCZs.

Size	150 m	180 m	210 m	240 m	270 m	300 m
City						
Harbin	0.463	0.587	0.583	0.845	0.613	0.590
Guangzhou	0.756	0.767	0.785	0.835	0.724	0.668
Xi'an	0.503	0.492	0.468	0.783	0.685	0.602

**Table A2.** Number of training and test samples for the 32 cities.

City	Sample	LCZ 1	LCZ 2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZ B	LCZD	LCZF	LCZG
Beijing	training	5,134	6,203	8,967	9,637	8,906	12,397	8,328	NA	8,257	2,815	10,529	3,672	6,131
	test	325	275	325	400	300	550	225	NA	325	250	325	225	550
Guangzhou	training	7,204	4,596	5,745	11,804	4,929	4,306	4,464	4,415	2,633	3,380	9,373	5,344	41,094
	test	300	275	250	425	300	300	450	100	250	300	1,325	300	1,500
Wuhan	training	1,627	2,373	3,170	5,585	9,362	4,459	22,345	NA	5,969	1,036	25,623	1,741	5,496
	test	150	250	200	325	375	325	675	NA	525	550	1,025	175	325
Shijiazhuang	training	3,812	5,859	11,715	9,161	9,889	3,991	7,685	4,931	2,472	1,635	18,428	3,135	4,403
	test	325	225	800	575	750	400	600	75	250	100	1,175	125	325
Lanzhou	training	1,380	2,204	2,650	3,730	3,260	NA	3,521	NA	3,695	2,885	5,468	10,873	1,932
	test	275	250	350	375	275	NA	325	NA	625	325	475	175	225
Hangzhou	training	2,363	2,288	4,497	5,564	5,146	8,197	19,844	NA	10,523	1,468	3,099	1,480	10,196
	test	175	225	200	275	300	350	300	NA	425	200	200	175	150
Harbin	training	2,829	2,611	4,688	6,367	5,503	8,051	6,986	NA	3,201	2,099	12,952	2,921	6,967
	test	350	325	425	500	400	425	600	NA	350	225	650	325	875
Changsha	training	2,013	2,856	2,864	7,937	4,523	6,794	6,161	NA	408	766	1,670	5,505	587
	test	125	200	200	250	175	200	325	NA	75	150	225	325	250
Zhengzhou	training	3,472	4,998	9,952	5,809	3,865	1,528	13,932	1,079	12,680	1,522	32,897	3,237	1,715
	test	200	275	375	300	175	75	400	125	325	175	525	250	150
Chongqing	training	4,218	459	3,269	6,227	7,596	NA	1,460	NA	4,302	2,749	15,899	54,377	9,910
	test	150	25	150	400	350	NA	225	NA	225	1,475	350	225	675
Chengdu	training	3,475	3,042	2,303	8,251	7,173	4,414	6,129	NA	4,780	3,021	2,244	8,799	4,826
	test	150	225	200	375	275	250	225	NA	150	200	175	200	225
Nanning	training	NA	4,393	4,804	5,406	4,940	1,940	5,273	NA	18,818	1,942	4,014	5,310	10,038
	test	NA	250	200	275	225	50	250	NA	625	200	275	175	350
Shenyang	training	6,705	6,622	5,053	6,977	7,715	7,382	6,111	NA	7,466	3,929	5,091	3,900	5,955
	test	350	325	375	300	350	350	700	NA	650	225	510	300	625
Tianjin	training	2,385	4,711	4,523	3,900	5,613	3,956	3,909	NA	5,138	2,514	4,547	6,647	5,034
	test	200	225	250	200	250	200	425	NA	525	250	560	300	150
Jinan	training	8,193	6,300	8,636	9,613	6,134	8,467	5,700	NA	5,078	8,382	8,727	6,731	9,766
	test	225	375	350	350	275	350	75	NA	200	175	585	275	150
Yinchuan	training	4,950	6,577	6,602	3,569	5,272	5,321	3,044	NA	559	50,680	31,438	4,204	7,898
	test	275	350	275	175	250	275	75	NA	150	575	925	525	650
Taiyuan	training	3,997	3,090	7,810	6,266	2,416	13,185	20,191	NA	18,801	1,027	13,647	3,315	2,977
	test	250	175	175	175	125	275	175	NA	400	125	550	250	175
Hohhot	training	5,710	7,681	7,396	2,465	5,298	3,863	9,134	NA	780	4,071	5,980	3,975	18,924
	test	350	275	300	150	150	275	175	NA	225	200	225	125	125
Nanjing	training	1,917	1,805	1,953	3,632	3,214	3,223	3,971	12,707	1,511	1,095	1,728	2,079	16,389
	test	275	200	225	375	225	375	500	200	175	250	175	100	475
Kunming	training	3,239	2,584	2,856	9,012	2,848	3,922	5,827	NA	52,115	4,755	12,652	15,737	3,000
	test	200	200	175	425	200	150	375	NA	800	325	250	525	350
Xining	training	3,398	5,187	8,112	1,260	914	NA	6,064	NA	NA	1,510	1,642	434	2,547
	test	175	225	275	75	50	NA	125	NA	NA	50	75	50	50
Lhasa	training	113	1,403	349	NA	NA	NA	NA	246	NA	1,656	NA	3,938	1,553
	test	25	75	25	NA	NA	NA	NA	25	NA	100	NA	175	75
Urumqi	training	NA	NA	1,855	3,394	6,077	NA	5,085	NA	NA	191	NA	65,448	8,700
	test	NA	NA	200	150	225	NA	175	NA	NA	75	NA	1,125	350
Haikou	training	NA	NA	733	2,220	1,909	486	1,395	NA	231	NA	348	152	875
	test	NA	NA	200	175	125	125	50	NA	75	NA	50	50	200
Taipei	training	1,567	1,444	1,564	NA	NA	NA	NA	NA	12,492	NA	325	NA	845
	test	100	75	175	NA	NA	NA	NA	NA	275	NA	75	NA	100

(Continued)

**Table A2.** (Continued).

City	Sample	LCZ 1	LCZ 2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZ B	LCZD	LCZF	LCZG
Fuzhou	training	2,212	1,786	2,408	5,332	2,387	2,312	4,083	1,535	1,217	718	3,358	5,133	10,645
	test	175	175	225	175	200	300	225	75	225	225	175	250	425
Guiyang	training	7,015	4,170	2,841	5,583	4,863	6,366	4,751	NA	3,911	5,032	4,362	3,717	5,595
	test	325	225	225	175	250	125	250	NA	200	250	175	250	275
Shanghai	training	1,823	3,384	4,070	10,813	6,882	7,060	5,455	10,720	2,775	1,969	11,467	1,513	6,405
	test	100	250	250	500	250	250	300	75	200	225	250	150	150
Nanchang	training	1,932	1,761	1,194	5,220	2,389	2,261	3,681	NA	544	2,033	353	2,669	1,265
	test	125	225	225	225	200	250	250	NA	200	225	200	250	150
Xian	training	3,495	3,485	9,408	13,770	5,455	2,609	8,069	NA	29,881	928	12,136	1,488	3,497
	test	225	200	300	325	300	200	350	NA	600	175	550	175	225
Changchun	training	2,868	4,320	5,005	9,117	10,166	6,588	13,618	NA	6,551	531	27,759	1,665	89,841
	test	225	225	250	375	300	225	525	NA	300	125	1,050	225	250
Hefei	training	2,212	1,786	2,408	5,332	2,387	2,312	4,083	1,535	1,217	718	3,358	5,133	10,645
	test	225	275	225	425	300	250	375	150	325	225	325	225	575

**Table A3.** Confusion matrix of the LCZ-CNN model for all 32 major cities of China.

	LCZ1	LCZ2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZB	LCZD	LCZF	LCZG	Total	PA
LCZ1	<b>4114</b>	527	41	1314	177	15	57	0	5	24	8	38	30	6350	0.65
LCZ2	171	<b>4849</b>	493	222	793	115	127	29	0	64	0	12	0	6875	0.71
LCZ3	102	485	<b>6903</b>	116	189	73	383	37	4	20	45	18	0	8375	0.82
LCZ4	918	118	78	<b>7150</b>	438	177	98	34	10	41	52	79	32	9225	0.78
LCZ5	162	573	75	1059	<b>5232</b>	549	56	30	11	95	33	21	29	7925	0.66
LCZ6	55	82	149	225	568	<b>5139</b>	158	52	30	248	91	75	28	6900	0.74
LCZ8	130	94	392	100	114	85	<b>8365</b>	30	78	110	25	171	31	9725	0.86
LCZ10	86	62	42	56	73	11	21	<b>469</b>	0	0	2	3	0	825	0.57
LCZA	40	10	75	41	49	24	117	1	<b>8891</b>	331	71	16	9	9675	0.92
LCZB	54	23	58	190	201	267	104	17	357	<b>5833</b>	452	315	79	7950	0.73
LCZD	30	21	61	48	44	131	65	4	87	386	<b>12006</b>	438	109	13430	0.89
LCZF	58	33	132	58	45	99	180	67	19	253	336	<b>6603</b>	117	8000	0.83
LCZG	246	0	14	31	81	35	48	12	20	47	104	353	<b>10134</b>	11125	0.91
Total	6166	6877	8513	10610	8004	6720	9779	782	9512	7452	13225	8142	10598	OA	0.81
UA	0.67	0.71	0.81	0.67	0.65	0.76	0.86	0.60	0.93	0.78	0.91	0.81	0.96	Kappa	0.79

**Table A4.** Confusion matrix of WUDAPT (Spec) for all 32 major cities of China.

	LCZ1	LCZ2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZB	LCZD	LCZF	LCZG	Total	PA
LCZ1	<b>3098</b>	829	85	1587	319	50	159	21	26	33	14	67	62	6350	0.49
LCZ2	380	<b>3811</b>	719	460	979	111	155	47	0	75	59	67	12	6875	0.55
LCZ3	127	713	<b>5975</b>	262	231	261	576	42	27	29	41	91	0	8375	0.71
LCZ4	1110	180	222	<b>5988</b>	954	276	128	59	28	105	30	104	41	9225	0.65
LCZ5	225	643	253	1557	<b>4217</b>	625	161	25	14	71	92	17	25	7925	0.53
LCZ6	105	104	268	611	674	<b>4069</b>	254	51	13	373	223	143	12	6900	0.59
LCZ8	136	159	372	226	129	58	<b>8131</b>	100	116	58	56	173	11	9725	0.84
LCZ10	137	39	61	92	45	18	60	<b>368</b>	0	4	0	1	0	825	0.45
LCZA	26	47	126	52	55	84	154	0	<b>8709</b>	157	200	8	57	9675	0.90
LCZB	66	78	186	419	343	553	154	16	371	<b>4463</b>	951	293	57	7950	0.56
LCZD	22	26	170	58	33	249	50	20	133	280	<b>10805</b>	607	977	13430	0.80
LCZF	158	13	288	118	131	303	471	97	22	188	483	<b>5201</b>	527	8000	0.65
LCZG	99	12	17	69	181	165	74	6	102	30	181	183	<b>10006</b>	11125	0.90
Total	5689	6654	8742	11499	8291	6822	10527	852	9561	5866	13135	6955	11787	OA	0.70
UA	0.54	0.57	0.68	0.52	0.51	0.60	0.77	0.43	0.91	0.76	0.82	0.75	0.85	Kappa	0.68

**Table A5.** Confusion matrix of WUDAPT (Spec+GLCM) for all 32 major cities of China.

	LCZ1	LCZ2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZB	LCZD	LCZF	LCZG	Total	PA
LCZ1	<b>3246</b>	770	50	1692	289	23	190	17	6	10	13	10	34	6350	0.51
LCZ2	269	<b>4015</b>	710	455	921	137	145	49	14	76	30	49	5	6875	0.58
LCZ3	95	610	<b>6066</b>	246	302	336	496	37	27	32	40	86	2	8375	0.72
LCZ4	1092	150	213	<b>6369</b>	760	231	114	31	3	113	27	89	33	9225	0.69
LCZ5	160	593	236	1476	<b>4456</b>	594	145	30	26	104	77	1	27	7925	0.56
LCZ6	67	115	180	634	666	<b>4293</b>	243	41	4	382	116	127	32	6900	0.62
LCZ8	116	123	316	188	95	72	<b>8276</b>	87	137	35	76	194	10	9725	0.85
LCZ10	68	88	50	164	19	5	41	<b>383</b>	0	4	1	2	0	825	0.46
LCZA	35	30	126	31	35	99	144	0	<b>8679</b>	174	194	41	87	9675	0.90
LCZB	48	53	161	343	276	564	180	31	431	<b>4406</b>	1092	293	72	7950	0.55
LCZD	23	33	125	100	71	233	59	0	116	363	<b>10884</b>	490	933	13430	0.81
LCZF	173	35	228	118	85	327	468	78	1	182	363	<b>5470</b>	472	8000	0.68
LCZG	73	6	38	99	156	184	57	0	155	82	168	166	<b>9941</b>	11125	0.89
Total	5465	6621	8499	11915	8131	7098	10558	784	9599	5963	13081	7018	11648	OA	0.72
UA	0.59	0.61	0.71	0.53	0.55	0.60	0.78	0.49	0.90	0.74	0.83	0.78	0.85	Kappa	0.69

**Table A6.** Confusion matrix of BoVW(Spec) for all 32 major cities of China.

	LCZ1	LCZ2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZB	LCZD	LCZF	LCZG	Total	PA
LCZ1	<b>3279</b>	744	152	1454	293	50	228	6	41	57	0	8	38	6350	0.52
LCZ2	379	<b>3995</b>	750	441	798	144	242	70	0	1	35	20	0	6875	0.58
LCZ3	77	588	<b>5969</b>	268	346	167	681	47	41	24	72	80	15	8375	0.71
LCZ4	988	264	135	<b>6331</b>	899	149	144	24	17	80	39	114	41	9225	0.69
LCZ5	102	873	208	1408	<b>4490</b>	439	163	17	9	82	123	10	1	7925	0.57
LCZ6	55	158	262	454	694	<b>4391</b>	255	42	63	175	211	112	28	6900	0.64
LCZ8	150	137	252	262	132	84	<b>7983</b>	60	135	52	147	303	28	9725	0.82
LCZ10	67	94	28	109	29	33	99	<b>271</b>	0	46	4	38	7	825	0.33
LCZA	25	11	73	34	48	119	106	0	<b>8653</b>	219	254	21	112	9675	0.89
LCZB	84	18	143	376	339	487	194	23	495	<b>4133</b>	1129	191	338	7950	0.52
LCZD	54	10	155	84	135	297	80	0	271	449	<b>10707</b>	556	632	13430	0.80
LCZF	182	30	367	123	148	256	717	52	35	260	531	<b>4985</b>	314	8000	0.62
LCZG	87	33	104	164	288	86	106	0	151	59	241	214	<b>9592</b>	11125	0.86
Total	5529	6955	8598	11508	8639	6702	10998	612	9911	5637	13493	6652	11146	OA	0.70
UA	0.59	0.57	0.69	0.55	0.52	0.66	0.73	0.44	0.87	0.73	0.79	0.75	0.86	Kappa	0.67

**Table A7.** Confusion matrix of BoVW (Spec+GLCM) for all 32 major cities of China.

	LCZ1	LCZ2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZB	LCZD	LCZF	LCZG	Total	PA
LCZ1	<b>3296</b>	771	146	1484	267	47	235	9	11	63	1	3	17	6350	0.52
LCZ2	397	<b>3916</b>	821	378	809	152	258	68	0	0	40	33	3	6875	0.57
LCZ3	79	563	<b>6129</b>	238	397	190	591	35	0	15	60	75	3	8375	0.73
LCZ4	907	236	131	<b>6517</b>	778	186	157	28	19	75	29	105	57	9225	0.71
LCZ5	95	824	198	1370	<b>4682</b>	436	147	15	7	91	51	9	0	7925	0.59
LCZ6	50	112	210	430	629	<b>4446</b>	316	49	63	169	186	213	27	6900	0.64
LCZ8	111	107	266	178	125	73	<b>8294</b>	47	98	74	113	215	24	9725	0.85
LCZ10	71	119	5	131	16	31	80	<b>323</b>	0	12	1	29	7	825	0.39
LCZA	1	15	50	52	29	102	56	0	<b>8907</b>	174	182	34	73	9675	0.92
LCZB	81	28	123	350	315	372	194	25	551	<b>4364</b>	984	132	431	7950	0.55
LCZD	49	28	91	121	89	267	103	0	223	518	<b>10471</b>	602	868	13430	0.78
LCZF	153	13	340	134	118	221	777	49	53	276	551	<b>4999</b>	316	8000	0.62
LCZG	82	40	116	223	265	79	118	0	91	194	181	86	<b>9650</b>	11125	0.87
Total	5372	6772	8626	11606	8519	6602	11326	648	10023	6025	12850	6535	11476	OA	0.71
UA	0.61	0.58	0.71	0.56	0.55	0.67	0.73	0.50	0.89	0.72	0.81	0.76	0.84	Kappa	0.69

**Table A8.** Confusion matrix of MLP for all 32 major cities of China.

	LCZ1	LCZ2	LCZ3	LCZ4	LCZ5	LCZ6	LCZ8	LCZ10	LCZA	LCZB	LCZD	LCZF	LCZG	Total	PA
LCZ1	<b>3532</b>	756	146	1254	395	54	86	15	0	26	0	11	75	6350	0.56
LCZ2	393	<b>4211</b>	650	404	818	105	207	26	1	37	11	12	0	6875	0.61
LCZ3	224	647	<b>6224</b>	210	235	145	465	29	7	44	5	140	0	8375	0.74
LCZ4	805	160	223	<b>6166</b>	1146	336	93	8	12	146	17	92	21	9225	0.67
LCZ5	169	573	285	1599	<b>4450</b>	557	70	5	32	149	21	7	8	7925	0.56
LCZ6	106	136	227	518	657	<b>4437</b>	205	31	49	280	83	160	11	6900	0.64
LCZ8	170	172	318	217	161	143	<b>8084</b>	42	64	92	23	204	35	9725	0.83
LCZ10	100	135	42	96	69	4	44	<b>324</b>	0	1	1	9	0	825	0.39
LCZA	56	27	64	51	60	89	109	2	<b>8726</b>	275	140	26	50	9675	0.90
LCZB	37	41	146	331	170	393	97	15	365	<b>5328</b>	675	288	64	7950	0.67
LCZD	4	10	150	75	87	135	70	0	134	563	<b>11663</b>	412	127	13430	0.87
LCZF	169	56	136	107	91	196	226	104	35	279	445	<b>5996</b>	160	8000	0.75
LCZG	124	21	28	111	351	45	57	6	66	145	114	229	<b>9828</b>	11125	0.88
Total	5889	6945	8639	11139	8690	6639	9813	607	9491	7365	13198	7586	10379	OA	0.74
UA	0.60	0.61	0.72	0.55	0.51	0.67	0.82	0.53	0.92	0.72	0.88	0.79	0.95	Kappa	0.72