Handbook of Geospatial Approaches to Sustainable Cities

Edited by Qihao Weng in collaboration with Cheolhee Yoo

First published 2024

ISBN: 978-1-032-15481-7 (hbk)
ISBN: 978-1-032-15534-0 (pbk)
ISBN: 978-1-003-24456-1 (ebk)

15

Comparisons of Deep Learning Models for Dynamic Local Climate Zone Mapping

Haojie Chen, Cheolhee Yoo, and Qihao Weng

DOI: 10.1201/9781003244561-19

The Open Access version of this chapter was funded by Qihao Weng.

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

15 Comparisons of Deep Learning Models for Dynamic Local Climate Zone Mapping

Haojie Chen, Cheolhee Yoo, and Qihao Weng

INTRODUCTION

The Local Climate Zone concept, which provides an objective protocol for measuring the magnitude of any city's climate and ecology, has been extensively used to address the limitation of using only urban and non-urban for urban climate and ecology issues (Xiang et al., 2023). The building structure and urban morphology were characterized to 10 built types based on building morphology factors and combing information on the surrounding exposure and surface conditions of a zone (Zhou et al., 2022). Therefore, this terminology has a great potential for revealing urban land changes through the portrayal of urban morphological characteristics. In addition, the analysis of multiple megacities using remote sensing images becomes feasible given the Earth Observation (EO)-based approach for classifying LCZ.

Local Climate Zone(LCZ) was designed from the very beginning with the purpose of providing an objective protocol for measuring the magnitude of the urban heat island effect in any city. With this term, it is not necessary to restrict the discussion of UHI(Urban Heat Island) to a distinction between urban and non-urban with just two options (Stewart & Oke, 2012). Standard set is divided into built types (1–10),and land cover types (A-G). Built types include categories 1–10, mainly distinguished by building height and impact. land cover types are also used as a means of differentiation, including paved and hard-packed (concrete, steel, stone, brick, tile, and concrete construction materials.), and permeable land cover types (low plants, scattered trees). Land cover types include categories A-G, which mainly distinguish between land cover landscape types, where the type of vegetation (height, species) is subdivided (Stewart, 2012).

There are four primary approaches to LCZ mapping, including the manual approach, GIS-based approach (Zhao, 2019), Earth Observation (EO)-based approach (Bechtel, 2012; Bechtel, 2015), and hybrid approach (Zhou, 2021). The EO-based WUDAPT method, for example, employs a standard workflow to produce LCZ maps using the same data source (landsat-8) and classifier (random forest), while

DOI: 10.1201/9781003244561-19

generating sample sets based on different regions (Bechtel, 2015). However, the accuracy of this method can be limited. In contrast, Convolutional Neural Networks (CNN) have demonstrated superior performance compared to other supervised deep learning algorithms in remote sensing visual recognition (Zhang, 2016; Ma, 2019). In the field of LCZ mapping, CNN models have been found to outperform traditional methods, such as random forests (Yoo, 2019).

This chapter delves into the potential of Convolutional Neural Networks (CNNs) in mapping Local Climate Zones (LCZs), providing a comparative study against traditional mapping methodologies. We introduce a novel data processing technique involving a 32*32 moving window and demonstrate its efficacy over conventional approaches. Leveraging data augmentation strategies for balanced LCZ representation, we evaluate the performance of several established CNN models in the context of LCZ mapping. With an overall accuracy of 80% to 90%, our findings highlight the promise of CNNs for future advancements in urban climate studies.

METHODOLOGY

There are four main approaches for LCZ mapping, the manual approach, GIS-based approach (C. Zhao, Jensen, Weng, Currit, & Weaver, 2019), EO-based (Earth Observation) approach (Bechtel et al., 2015; Bechtel & Daneke, 2012) and hybrid approach (Zhou et al., 2021). The manual method requires a lot of human and financial resources, not suitable for large scale mapping needs, it is even less possible to analyze multiple cities.

GIS-Based Approach

The basis for the GIS-based approach to map LCZ is based on values of geometric and surface cover properties, including sky view factor(SKF), aspect ratio(AR), building surface fraction(BSF), impervious surface fraction(ISF), pervious surface fraction(PSF), height of roughness elements(Height), and terrain roughness class (Stewart & Oke, 2012). Different research scholars have described these indices based on different data sources methods. Basic input data includes satellite and aerial imagery, land cover datasets, data with height information, road/street maps, building footprint maps and others (Oliveira et al., 2020; C. Zhao et al., 2019). The most used methods are the transformation based on existing land use datasets (Oliveira et al., 2020; C. Zhao et al., 2019) and other standard datasets for reclassification (Hidalgo et al., 2019)，and rule-based fuzzy preliminary classification using linear interpolation scores or ranks based on a representative set of metrics at the site, grid, subdistrict or block level (Kotharkar & Bagade, 2018; Lelovics et al., 2014; Quan, 2019). GIS-based methods are often inseparable from data with elevation information (e.g., NDSM or 3D city models) because elevation data are directly or indirectly related to SKF, ISF, BSF, PSF, and height of roughness elements (C. Zhao et al., 2019). The open source DSM and DEM data are both too coarse in spatial resolution, and additionally cannot be analyzed over multiple times

(Zhou et al., 2021). Precise point clouds can be obtained from LIDAR data to obtain NDSM elevation information, but due to its expensive equipment and cumbersome processing procedures, it is often used for small-scale measurements and usage. Although some cities or tracts have obtained airborne LIDAR range data, it is still difficult to obtain for the rest of the cities, and it is even more difficult to obtain at different times (C. Zhao et al., 2019). The disadvantage of using the LULC data for conversion to LCZ is that the definition of land use data varies from place to place, and there is no way to develop a uniform method for conversion (C. Zhao et al., 2019).

EO-Based Approaches

EO-based approaches currently use machine learning methods for classification, which are classified as Object-based approach and pixel-based approach. Object-based classification requires a lot of operations, more detailed information, and contextual knowledge (Ma et al., 2021a; Ma et al., 2021b), while the pixel-based approach can ignore some details and classify use pixel blocks, which is widely used in LCZ classification (Bechtel & Daneke, 2012). The approach involves selecting and solving the feature vector (to describe the most accurate and comprehensive information) in advance based on the features of the EO observations (to classify the LCZ), and then using classifiers to judge the feature vector, the main classifiers being random forests, support vector machines, and so on (Bechtel et al., 2015; Bechtel & Daneke, 2012). The method of selecting the different characteristics of Earth observation data and their best fit with LCZ observations has been studied by research scholars (Bechtel & Daneke, 2012). The WUDAPT method has developed a standard workflow for producing LCZ maps, which uses the same data source (landsat-8) and uses the same classifier (random forest) but can produce sample sets based on different regions, thus allowing scholars to produce LCZ maps that meet their needs based on their own study area (Bechtel et al., 2015). In contrast to deep learning, the WUDAPT method of shallow learning is used by a wide range of researchers because it is computationally convenient, takes less time and is less complex. However, this method is often not very accurate (R. Wang et al., 2018). Hence, many researchers have attempted to map the LCZ using multi-source Earth observation data sources, such as sentinel-1 (La, Bagan, & Yamagata, 2020) and sentinel-2 data (Shi & Ling, 2021), NTL data (visible light emerging from the Earth at night) (La et al., 2020), and so on. The multi-source data make the accuracy improved, but the ease of data acquisition raises the difficulty.

Hybrid Models

The hybrid model combines the advantages and disadvantages of both GIS-based and EO-based approaches. It has been verified by scholars that the accuracy of the LCZ maps obtained by the GIS-based method is generally higher than that of the WUDAPT method (R. Wang et al., 2018). The main reason is that the GIS-based

approach is able to operate strictly according to the rules of the LCZ land cover attributes. And because of the different data sources, the LCZ blocks are represented in different forms (e.g., vector data and grid areas), which enables to obtain sufficient size and boundaries for each LCZ class (Kotharkar & Bagade, 2018; Oliveira et al., 2020). The biggest drawback of the GIS-based approach is that data sources vary in quality and format, and even the methods of conversion are developed based on specific data sources (Quan, 2019). Some of the data are not global and data acquisition becomes difficult if we change the study area. Although the MUDAPT method is very simple and convenient for more research scholars to be used in many research areas, the disadvantage is that the method of WUDAPT is a supervised learning method, which requires manual judgment of the class of training samples in advance, and the quality of label data very much affect the final results. In a grid block, different types of land use and land cover types are often mixed due to the dense characteristics of the city, and the criteria for labeling by different scholars are also different (Quan, 2019).Some scholars, in order to synthesize the advantages and disadvantages of the above two methods, proposed to use EO observation data for LCZ classification while combining GIS data such as OSM data to help segment LCZ blocks, which has greatly improved the final classification effect (Fonte et al., 2019; Shi & Ling, 2021; R. Wang et al., 2018; Zhou et al., 2021).

Convolutional Neural Networks (CNN) Models

Convolutional neural networks (CNN) have been shown to outperform most deep learning supervised algorithms in the field of remote sensing visual recognition (Ma et al., 2019; Zhang et al., 2016), such as scene classification (Hu et al., 2015; Scott et al., 2017; Yuan, Han, Gu, & Yan, 2021), target detection (Shin et al., 2016), semantic segmentation (Cao et al., 2019). With its powerful ability to extract highly abstracted, generalized features, CNN models have been tried in the field of LCZ mapping and proven to be better than methods such as random forests (Yoo et al., 2019). So2LCZ42 provides a standardized dataset of 42 city clusters worldwide (plus 10 smaller regions), resulting in approximately 500,000 pairs of Sentinel-1 and Sentinel-2 image block data, and a carefully designed evaluation process resulting in an overall confidence level of 85% (Zhu et al., 2020). Several studies have shown that training with the So2Sat LCZ42 dataset can be trained to obtain better classification models (Chunping, Schmitt, Lichao, & Xiaoxiang; Qiu et al., 2019; Qiu et al., 2020; Zhou et al., 2021).

Most of the CNN models used for LCZ classification are adaptations of classical models, such as VGG, ResNet, etc. (X. Huang, Liu, & Li, 2021; S. Liu & Shi, 2020; Qiu et al., 2019; Qiu et al., 2020; Yoo et al., 2019; Zhou et al., 2021), while most of the models are trained from the initialized parameters. It is undeniable that none of us can prove the superiority of the models designed by the authors.

Some scholars are beginning to recognize that completely designing and training CNN models from scratch is very challenging because it requires large amounts of data and takes a long time (Yuan et al., 2021), and that CNNs

pre-trained on large datasets have general feature extraction capabilities that can be transferred to regions with smaller amounts of data (Shin et al., 2016). Many scholars have applied migratory learning for scene classification, and they have explored the effects of different depth levels of features on migratory learning, as well as the effects of different learning modes (parameter freezing, fine-tuning, initialization from random) on migratory learning results (Hu et al., 2015; Pires De Lima & Marfurt, 2019; H. Zhao, Liu, Zhang, & Liang, 2019). However, no scholars have tested in detail the best patterns and classifiers for migration learning for LCZ classification.

In summary, the CNNs-based approach is able to obtain a more generalized LCZ classification model by training. This will be far different from previous GIS-based and shallow learning-based methods, as GIS-based methods usually have a small number of fixed-time data sources and shallow learning-based methods usually have a small number of training samples. In particular, CNNs-based methods based on migration learning can lead to a very generalized model, which makes it possible to study the temporal variation of LCZs at the urban scale. However, the CNNs-based approach still suffers from low accuracy and a discrete distribution of classification results.

In this paper, a variety of CNNs models are tested using a migration learning approach in order to obtain a more generalized model that can enable the study of the temporal variation of LCZs at the urban scale. At the same time, experiments are carried out to obtain high accuracy CNNs models.

Data Processing and CNN Models Training

In our scheme, we considered the surrounding areas around the focal pixel, using a moving window of size 32*32 at 32-pixel intervals. This approach resulted in better results than the RF-based scheme and other CNN-based schemes (e.g., 10*10 size feature extraction block) (Yoo, 2019). To ensure an equivalent number of polygons and patches of different LCZ types and independent LCZ types from each other, we utilized data augmentation in each polygon (i.e., by rotating 90°, 180°, and 270°). Polygons refer to manually annotated polygonal regions that are homogeneous in nature, and batches represent a 32*32-pixel block. A list of the training and test datasets of LCZ types for each city is provided in Table 15.1.

This study testes the traditional CNN models, including VGG (Simonyan, 2014), ResNet (He, 2016), DenseNet (Huang, 2017), and GoogLeNet (Szegedy, 2014) to obtain accurate LCZ maps of the four megacities and synchronously perform the massive LCZ classification more swiftly. All initial parameters were sourced from the official pre-training parameters, the learning rates were set to 0.001, and the entire training lasted 100 epochs.

This study reclassified the LCZ types when calculating the overall accuracy (OA) to obtain highly precise results in classifying the three urbanization processes. OA_{ex}, OA_{high}, and $OA_{compact}$ were calculated to indicate the OA of the three urbanization processes, namely, new urbanization, intensified height, and intensified compactness, respectively.

$$OA = \begin{cases} OA_{ex} = N_{ex}/N_m \\ OA_{high} = N_{high}/N_m \\ OA_{compact} = N_{compact}/N_m \end{cases} \quad (15.1)$$

where N_{ex}, N_{high}, and $N_{compact}$ indicate the numbers of accurately judged samples for LCZ type of new urbanization, intensified height, and intensified compactness, respectively. N_m expresses the total number of all samples.

Since its introduction, convolutional neural networks (CNNs) have been widely used in image classification due to their powerful ability in extracting abstract features (Zhang et al., 2016), and many related models have been proposed and performed well in the standard dataset ImageNet, including VGG, ResNet, SqueezeNet, DenseNet, and GoogleNet. Each model has its own characteristics.

TABLE 15.1
Training and Test Dataset of Each LCZ Type by City

LCZ	Beijing			
	2020		2000	
	Training	Test	Training	Test
1	2* (456)	2* (234)	–	
2	21 (1344)	6 (696)	19 (918)	7 (372)
3	7 (498)	2 (138)	13 (1488)	1 (564)
4	33 (2184)	22 (768)	12 (270)	5 (132)
5	3* (48)	3* (24)	4 (228)	2 (48)
6	13 (834)	5 (402)	3* (96)	3* (72)
7	13 (684)	11 (270)	22 (1134)	6 (492)
8	2* (30)	2* (18)	–	
9	–		–	
10	13 (1638)	4 (1260)	5 (156)	1 (132)
A	3 (2058)	6 (840)	6 (1650)	1 (600)
B	3 (60)	1 (30)	–	
C	–		–	
D	18 (3132)	12 (1278)	16 (3060)	6 (1482)
E	3 (2592)	3 (1440)	3* (726)	3* (330)
F	–		7 (582)	3 (228)
G	7 (834)	3 (276)	8 (474)	3 (258)

* (The red-star classes) has been used to refer to situations with only a few reference polygons of the LCZ classes (no more than 3) (Yoo, 2019). The numbers in the table represent the number of polygons, whereas the figures in parentheses represent the number of patches.

VGG

VGG is one of the most classic neural network models, using small convolutional kernels (3*3) to extract more complex information than the previous large convolutional kernels (5*5 and 7*7). VGG consists of multiple blocks, each consisting of two or three convolutional layers, and a down sampling or pooling layer, which allows features of different depths to be extracted. The feature extraction layers are followed by three fully connected layers to classify the final result (Simonyan & Zisserman, 2014). In this paper, the input layers (3*225*225 to 7*32*32) are modified for the application. VGG is mainly divided into VGG13, VGG16, and VGG19 according to the depth of the network.

ResNet

ResNet uses a residual module to enable the network to be trained more deeply. Through the VGG model scholars found that deeper networks can extract deeper features and information, but if too much training will also appear overfitting phenomenon, in order to avoid this phenomenon, ResNet adds a residual module so that the gradient can be passed between the front and back layers of the network (He et al., 2016). In this paper, the input channels are changed to seven channels. ResNet is divided into ResNet18, ResNet34, ResNet50, ResNet101, and ResNet154.

SqueezeNet

SqueezeNet uses the Fire module to encapsulate multiple convolutional and down sampling layers, consisting of Squeeze (1*1 convolutional kernel) and Expand (1*1 convolutional kernel and 3*3 convolutional kernel) (Iandola et al., 2016). In this paper, the input channels are changed to seven channels. SqueezeNet includes SqueezeNet 1.0 and SqueezeNet 1.1 versions.

DenseNet

As mentioned earlier, in CNNs, the gradient disappearance and overfitting problems become more and more obvious as the depth of the network deepens. ResNet proposes a residual module that enables the fusion of the front and back gradients within the module, but cannot fuse features that span large distances. DenseNet continues the idea of feature fusion by fusing the features of all modules and trying to ensure that layer-to-layer DenseNet continues the idea of feature fusion by fusing the features of all modules and trying to ensure the maximum continuity of information between layers (G. Huang, Liu, Van Der Maaten, & Weinberger, 2017). In this paper, the input channels are changed to seven channels. DenseNet is mainly divided into DenseNet121, DenseNet161, DenseNet169,and DenseNet201.

GoogLeNet

GoogLeNet is composed of several inception modules, and the inception modules internally fuse features from multiple convolutional kernels (1*1, 3*3, 5*5), and this approach enables the fusion of different scale features (Szegedy et al., 2014). At the same time, the deeper the network is, the larger the proportion of 3*3 and 5*5 is, the more abstract features can be extracted. In this paper, the input channels are changed to seven channels.

LCZ MAPPING RESULTS

The performance evaluation of various Convolutional Neural Network (CNN) models for Beijing is presented in Table 15.2, case of LCZ classification results for the Beijing region are shown in Figure 15.1. The models include DenseNet121,

TABLE 15.2
The Performance Evaluation of the CNN Models for Beijing, Demonstrating High Overall Accuracy (OA) Ranging from 80% to 90%

Country	Year	CNNs	OA_{all}	OA_{ex}	OA_{high}	$OA_{compact}$
		DenseNet121	0.8243	0.9906	0.9450	0.9252
		GoogleNet	0.8603	0.9867	0.9432	0.9134
		ResNet50	0.8420	0.9886	0.9289	0.9216
		VGG16	0.8450	0.9864	0.9297	0.9097
		VGG19	0.8324	0.9853	0.9222	0.9038
		DenseNet121	0.8529	0.9630	0.9550	0.9059
		GoogleNet	0.8548	0.9655	0.9618	0.9277
		ResNet50	0.7208	0.9558	0.9459	0.8960
		VGG16	0.8047	0.9675	0.9448	0.9302
		VGG19	0.7295	0.9456	0.9463	0.8714

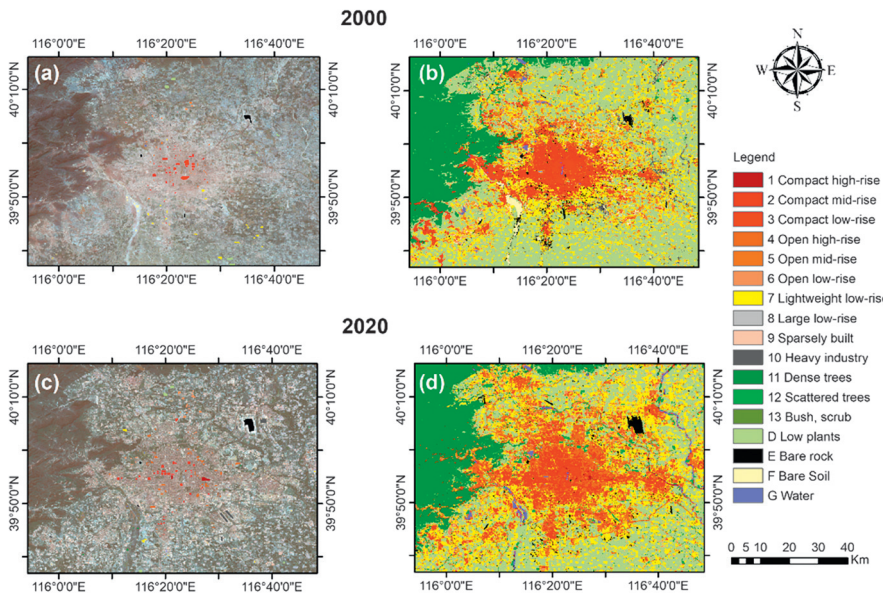

FIGURE 15.1 LCZ classification results for Beijing. (a) and (c) indicate the references of 2000 and 2020, respectively; (b) and (d) indicate the LCZ classification results of 2000 and 2020, respectively.

GoogleNet, ResNet50, VGG16, and VGG19. These models demonstrate high overall accuracy (OA), with values ranging from 80% to 90%. The DenseNet121 model achieved an accuracy of 82.43% in the first instance and improved to 85.29% in the second instance. GoogleNet's performance was slightly better, with accuracy scores of 86.03% and 85.48% in the first and second instances, respectively. The ResNet50 model's performance was consistent, with accuracy scores of 84.20% and 72.08% in the first and second instances, respectively. The VGG16 model showed a decrease in performance from the first instance (84.50% accuracy) to the second instance (80.47% accuracy). Similarly, the VGG19 model's performance decreased from 83.24% accuracy in the first instance to 72.95% accuracy in the second instance. The performance results across the models suggest that GoogleNet and DenseNet121 provided the most reliable and consistent results in this assessment.

DISCUSSION AND CONCLUSION

In our research, we examined the performance of various classical convolutional neural network (CNN) models, including VGG, ResNet, DenseNet, and GoogLeNet, in local climate zone (LCZ) classification tasks. In our experiments, we found that DenseNet and GoogLeNet provided the most reliable and stable results in our tasks. This could be because these models effectively integrate features of various scales and ensure the maximum continuity of information between the layers of the network as much as possible. Particularly, DenseNet, by integrating the features of all modules and trying to ensure the maximum continuity of information between the network layers, solves the problems of gradient disappearance and overfitting in deep networks.

Compared to that, although VGG and ResNet also showed high accuracy in some situations, their performance varied between the two instances, suggesting that they may not have the stability of DenseNet and GoogLeNet (Huang, 2021; Li, 2021). However, it is important to note that although these models perform well in our tasks, their performance may vary in other tasks (Yu, 2022). Therefore, we recommend researchers to choose the most suitable model based on their specific tasks and data.

In summary, our research shows that by using suitable CNN models, we can effectively classify and map local climate zones (LCZ), and further understand the process of urbanization. However, to optimize performance and accuracy, it is crucial to choose the most suitable model and tune parameters. Future work might explore more models and methods to further improve the accuracy and efficiency of LCZ classification (Liang, 2023).

REFERENCES

Bechtel, B., Alexander, P., Böhner, J., Ching, J., Conrad, O., Feddema, J., … Stewart, I. (2015). Mapping local climate zones for a worldwide database of the form and function of cities. *ISPRS International Journal of Geo-Information*, 4(1), 199–219. doi: 10.3390/ijgi4010199

Bechtel, B., & Daneke, C. (2012). Classification of local climate zones based on multiple earth observation data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4), 1191–1202. doi: 10.1109/jstars.2012.2189873

Cao, Z., Fu, K., Lu, X., Diao, W., Sun, H., Yan, M., ... & Sun, X. (2019). End-to-end DSM fusion networks for semantic segmentation in high-resolution aerial images. *IEEE Geoscience and Remote Sensing Letters*, 16(11), 1766–1770.

Fonte, C. C., Lopes, P., See, L., & Bechtel, B. (2019). Using OpenStreetMap (OSM) to enhance the classification of local climate zones in the framework of WUDAPT. *Urban Climate*, 28, 100456. doi:10.1016/j.uclim.2019.100456.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Hidalgo, J., Dumas, G., Masson, V., Petit, G., Bechtel, B., Bocher, E., … Mills, G. (2019). Comparison between local climate zones maps derived from administrative datasets and satellite observations. *Urban Climate*, 27, 64–89. doi:10.1016/j.uclim.2018.10.004.

Hu, F., Xia, G.-S., Hu, J., & Zhang, L. (2015). Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. *Remote Sensing*, 7(11), 14680–14707. doi:10.3390/rs71114680.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Huang, X., Chen, W., & Yang, W. (2021). Improved algorithm based on the deep integration of googlenet and residual neural network. Paper presented at the Journal of Physics: Conference Series.

Huang, X., Liu, A., & Li, J. (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.

Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv preprint arXiv:1602.07360*

Kotharkar, R., & Bagade, A. (2018). Local Climate Zone classification for Indian cities: A case study of Nagpur. *Urban Climate*, 24, 369–392. doi:10.1016/j.uclim.2017.03.003.

La, Y., Bagan, H., & Yamagata, Y. (2020). Urban land cover mapping under the Local Climate Zone scheme using Sentinel-2 and PALSAR-2 data. *Urban Climate*, 33, 100661. doi:10.1016/j.uclim.2020.100661.

Lelovics, E., Unger, J., Gál, T., & Gál, C. (2014). Design of an urban monitoring network based on Local Climate Zone mapping and temperature pattern modelling. *Climate Research*, 60(1), 51–62. doi:10.3354/cr01220.

Li, Q., Yang, Y., Guo, Y., Li, W., Liu, Y., Liu, H., & Kang, Y. J. I. A. (2021). Performance Evaluation of Deep Learning Classification Network for Image Features. 9, 9318–9333.

Liang, Y., Song, W., Cao, S., & Du, M. J. R. S. (2023). Local Climate Zone Classification Using Daytime Zhuhai-1 Hyperspectral Imagery and Nighttime Light Data. 15(13), 3351.

Liu, S., & Shi, Q. (2020). Local climate zone mapping as remote sensing scene classification using deep learning: A case study of metropolitan China. *ISPRS Journal of Photogrammetry and Remote Sensing*, 164, 229–242. doi:10.1016/j.isprsjprs.2020.04.008.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166–177. doi:10.1016/j.isprsjprs.2019.04.015

Ma, L., Yang, Z., Zhou, L., Lu, H., & Yin, G. (2021a). Local climate zones mapping using object-based image analysis and validation of its effectiveness through urban surface temperature analysis in China. *Building and Environment*, 206, 108348. doi:10.1016/j.buildenv.2021.108348.

Ma, L., Zhu, X., Qiu, C., Blaschke, T., & Li, M. (2021b). Advances of Local Climate Zone Mapping and Its Practice Using Object -Based Image Analysis. *Atmosphere*, 12(9), 1146. doi:10.3390/atmos12091146.

Oliveira, A., Lopes, A., & Niza, S. (2020). Local climate zones in five southern European cities: An improved GIS-based classification method based on Copernicus data. *Urban Climate*, 33, 100631. doi:10.1016/j.uclim.2020.100631.

Pires De Lima, R., & Marfurt, K. (2019). Convolutional Neural Network for Remote-Sensing Scene Classification: Transfer Learning Analysis. *Remote Sensing*, 12(1), 86. doi:10.3390/rs12010086.

Qiu, C., Mou, L., Schmitt, M., & Zhu, X. X. (2019). Local climate zone-based urban land cover classification from multi-season alSentinel-2 images with a recurrent residual network. *ISPRS Journal of Photogrammetry and Remote Sensing*, 154, 151–162. doi:10.1016/j.isprsjprs.2019.05.004.

Qiu, C., Tong, X., Schmitt, M., Bechtel, B., & Zhu, X. X. (2020). Multilevel Feature Fusion-Based CNN for Local Climate Zone Classification From Sentinel-2 Images: Benchmark Results on the So2Sat LCZ42 Dataset. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 2793–2806. doi:10.1109/jstars.2020.2995711.

Quan, J. (2019). Enhanced geographic information system-based mapping of local climate zones in Beijing, China. *Science China Technological Sciences*, 62(12), 2243–2260. doi:10.1007/s11431-018-9417-6.

Scott, G. J., England, M. R., Starms, W. A., Marcum, R. A., & Davis, C. H. (2017). Training Deep Convolutional Neural Networks for Land–Cover Classification of High-Resolution Imagery. *IEEE Geoscience and Remote Sensing Letters*, 14(4), 549–553. doi:10.1109/lgrs.2017.2657778.

Shi, L., & Ling, F. (2021). Local Climate Zone Mapping Using Multi-Source Free Available Datasets on Google Earth Engine Platform. *Land*, 10(5), 454. doi:10.3390/land10050454.

Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., . . . Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298. doi:10.1109/tmi.2016.2528162.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12), 1879–1900. doi:10.1175/bams-d-11-00019.1

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., …Rabinovich, A. (2014). Going Deeper with Convolutions. arXiv pre-print server. doi: Nonearxiv:1409.4842

Wang, R., Ren, C., Xu, Y., Lau, K. K.-L., & Shi, Y. (2018). Mapping the local climate zones of urban areas by GIS-based and WUDAPT methods: A case study of Hong Kong. *Urban Climate*, 24, 567–576. doi:10.1016/j.uclim.2017.10.001.

Xiang, Y., Cen, Q., Peng, C., Huang, C., Wu, C., Teng, M., & Zhou, Z. (2023). Surface urban heat island mitigation network construction utilizing source-sink theory and local climate zones. *Building and Environment*, 243, 110717.

Yoo, C., Han, D., Im, J., & Bechtel, B. (2019). Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas

using Landsat images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 157, 155–170. doi:10.1016/j.isprsjprs.2019.09.009

Yu, Z., Dong, Y., Cheng, J., Sun, M., Su, F. J. S., & Networks, C. (2022). Research on Face Recognition Classification Based on Improved GoogleNet. 2022, 1–6.

Yuan, B., Han, L., Gu, X., & Yan, H. (2021). Multi-deep features fusion for high-resolution remote sensing image scene classification. *Neural Computing and Applications*, 33(6), 2047–2063. doi:10.1007/s00521-020-05071-7.

Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2), 22–40. doi:10.1109/mgrs.2016.2540798

Zhao, C., Jensen, J., Weng, Q., Currit, N., & Weaver, R. (2019). Application of airborne remote sensing data on mapping local climate zones: Cases of three metropolitan areas of Texas, U.S. *Computers, Environment and Urban Systems*, 74, 175–193. doi:10.1016/j.compenvurbsys.2018.11.002

Zhao, H., Liu, F., Zhang, H., & Liang, Z. (2019). Convolutional neural network based heterogeneous transfer learning for remote-sensing scene classification. *International Journal of Remote Sensing*, 40(22), 8506–8527. doi:10.1080/01431161.2019.1615652.

Zhou, L., Yuan, B., Hu, F., Wei, C., Dang, X., & Sun, D. (2022). Understanding the effects of 2D/3D urban morphology on land surface temperature based on local climate zones. *Building and Environment*, 208, 108578.

Zhou, Y., Wei, T., Zhu, X., & Collin, M. (2021). A parcel-based deep-learning classification to map local climate zones from sentinel-2 images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 4194–4204. doi:10.1109/jstars.2021.3071577

Zhu, X. X., Hu, J., Qiu, C., Shi, Y., Kang, J., Mou, L., . . . Wang, Y. (2020). So2Sat LC Z42: A Benchmark Data Set for the Classification of Global Local Climate Zones [Software and Data Sets]. *IEEE Geoscience and Remote Sensing Magazine*, 8(3), 76–89. doi:10.1109/mgrs.2020.2964708.