MULTI-LABEL LOCAL CLIMATE ZONE MAPPING AS SCENE CLASSIFICATION USING VERY HIGH RESOLUTION IMAGERY: PRELIMINARY RESULT OF HONG KONG

Shengjie Liu[1], *Qian Shi*[2]

[1]Department of Physics, The University of Hong Kong, Pokfulam, Hong Kong
[2]School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China

ABSTRACT

With the near completion of WUDAPT Level 0 data, one of the next goals is to generate more accurate and detailed local climate zone (LCZ) maps. An important issue is how to integrate building height information into LCZ maps. We here present a multi-label classification method using very high resolution (VHR) imagery to implicitly integrate building height information. Since we humans can tell whether a place is high-rise or not based on the shading of buildings and the surrounding context, it is possible to extract such information using deep learning methods. We use Hong Kong as case study and show the potential of LCZ mapping with VHR imagery in distinguishing small-scale landscape features like city parks. The multi-label LCZ maps also provide a solution to generate fine-grained subclass LCZ mapping, in which a place can be classified as a combination of multiple LCZs, e.g., compact low-rise with open high-rise.

Index Terms— Multi-label classification, local climate zone, mapping, very high resolution, subclass

1. INTRODUCTION

Fine-grained mapping of local climate zone (LCZ) is a next goal of the World Urban Database and Access Portal Tools (WUDAPT) project. Currently, LCZ maps are generated into 17 classes. The common protocol is to translate free-of-charge medium-resolution satellite images into LCZ maps. Among them, Landsat with 30 m spatial resolution and Sentinel-2 with 10 m resolution are the most often used image sources [1, 2]. Although these medium-resolution images achieve satisfactory results in generating current LCZ maps, they lack the contextual details for more complex LCZ mapping, e.g., with building height considered [3].

Building height is an important part in the next generation of WUDAPT, which should significantly distinguish high-rise buildings from the low-rise ones. However, height estimation by Sentinel-1 PolSAR data has suffered in LCZ mapping. In many cases, the integration of Sentinel-1 PolSAR data did not help and even degraded the mapping accuracy [4].

New advanced machine learning models now can estimate building height from a single very high resolution (VHR) image due to the shading of buildings. Height estimation from a single image was also a track of the 2019 IEEE GRSS Data Fusion Contest [5]. Thus, it should be possible to extract and integrate height information for LCZ mapping. With the height information and the increase of spatial resolution, VHR images are promising in generating more accurate urban climate maps.

An extra benefit is the ability to further distinguish details of an urban place. Kotharkar and Bagade (2018) [6] show that the subclass LCZ mapping can represent a more complex combination of the natural and built environment. They generate an LCZ map for the Nagpur city in India with subclass LCZs such as compact mid-/low-rise (LCZ-2/3) and compact/light-weight low-rise (LCZ-3/7). As pointed out by them, the complex natural and built environment requires a more detailed classification of urban climate zones. This is especially the case for Asian cities, since the land use pattern of Asian cities is more mixed and fragmented [7].

Therefore, in this study, we use Hong Kong as a case to show the potential of multi-label subclass LCZ mapping in a complex high-rise high-density city. Its relatively small-scale urban design and high spatial heterogeneity make it a perfect case for detailed LCZ mapping [8].

2. DATA AND METHOD

2.1. VHR Imagery for Multi-label LCZ

We constructed a multi-label LCZ dataset for this study. The dataset consists of 72,494 VHR images from Google and Bing satellite images in the Pearl River Delta region, in which images collected in Hong Kong account for 18.3%, as shown in Table 1. These VHR images were collected at the zoom level of 18 with a spatial resolution of better than 1 m (close to 0.6 m depending on the latitude). Each image tile is in a size of 256×256.

Using web map satellite imagery to create LCZ maps allows us to easily integrate other information for accurate clas-

This work was supported by the National Natural Science Foundation of China under Grant 61976234. Contact: S. Liu (sjliu.me@gmail.com); Q. Shi (shixi5@mail.sysu.edu.cn).

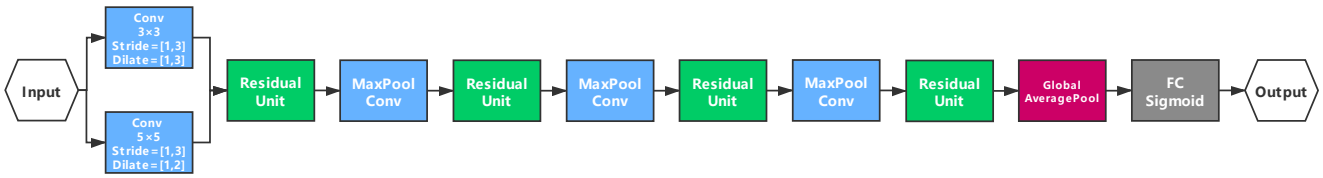


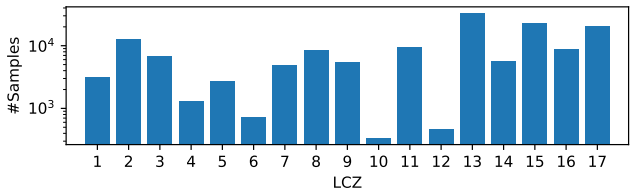
Fig. 1: Network used in this study.

Table 1: Number of labeled images of the dataset

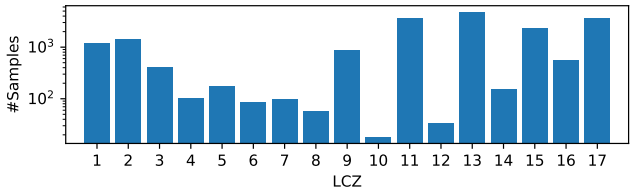
City	Number of Images	Percentage (%)
Guangzhou	13596	18.8
Shenzhen	13203	18.2
Foshan	4644	6.4
Dongguan	4579	6.3
Huizhou	4493	6.2
Zhongshan and Jiangmen	9514	13.1
Zhuhai and Macau	9209	12.7
Hong Kong	13256	18.3
Total	72494	100

sification, e.g., road network density. It also provides a natural hierarchy in generating multiple scales of LCZ maps, as each image tile can be divided into four sub-image tiles in the next zoom level. The tile system of popular online web maps uses the Spherical Web Mercator projection [9]. At a zoom level of 18, the spatial resolution is roughly 0.6 m at the Equator. For a 256×256 image tile, an LCZ covers roughly a size of 153×153 m^2.

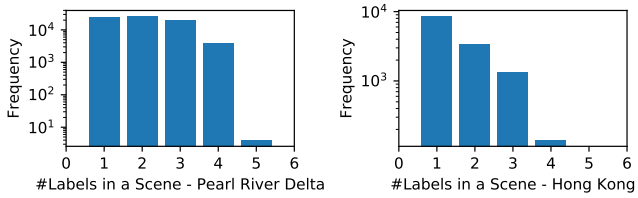
In the annotation process, each image is labeled independently based on its own urban context (the 153×153 m^2 area). To reduce the output space, we restricted the number of classes up to three, though in some extreme cases four or five was allowed, as shown in Fig. 2c. The class distribution is imbalance as shown in Figs. 2a and 2b (note the y-axis is in log scale). LCZ-13 bush is the most popular while LCZ-10 heavy industry is the least popular. The annotation was processed by several experts and then cross-validated once to check for incorrectness. But we should still be careful in using these datasets. As pointed out by [10], we should pay attention to the label quality even for the popular ImageNet dataset.



(a) Distribution of annotated LCZ classes (Pearl River Delta)



(b) Distribution of annotated LCZ classes (Hong Kong)



(c) Distribution of number of labels in a scene.

Fig. 2: Summary of statistics of the multi-label dataset.

2.2. Multi-label convolutional neural network

We design a multi-scale convolutional neural network (CNN) as a baseline for multi-label LCZ mapping, as shown in Fig. 1. It consists of four residual units [11], and we replace the last function from softmax to sigmoid to achieve multi-label classification. A total of three settings are being tested. The first one uses stride=1 and dilated rate=1 (Standard). The second one uses stride=3 and dilated rate=1 (Light-weight). The third one uses stride=1 and dilated rate=3. A high stride reduces computation by reducing the size of feature maps, which may be useful in a 256×256 image classification task. A high dilated rate empowers convolutions to shave a large receptive field.

2.3. Evaluation metrics

The evaluation of multi-label classification can be tricky, as the output space becomes exponentially large due to the possible combinations of coexist labels [12]. Here, we select the Hamming loss (hloss), ranking loss (rankloss), instance accuracy (acc) and F_β metrics to examine the multi-label classification performance. We set the β value as one. The two loss metrics are the smaller the better, and the acc and F_β are the larger the better. For details of these evaluation metrics, we refer readers to the review paper by Wu and Zhou (2017) [12].

3. RESULTS AND ANALYSIS

Experiments are conducted with TensorFlow 1.9. We use the AdaDelta optimizer [13] with a learning rate of 1.0 for 80 epochs and of 0.1 for another 20 epochs to train the network. The reported results are averaged from five random runs with a fixed train-test split: half for training and half for testing.

3.1. Single-Label VHR LCZ Maps

To obtain single-label LCZ maps via the multi-label dataset, we adopt a simple solution to treat the class with the maximum classification probability as the dominant class. The map can be found at https://sjliu.me/lczmap/hongkong, in which you can check the LCZ map with satellite image simultaneously.

We here show some samples of downtown Hong Kong in Fig. 3. The LCZ map via VHR imagery captures the heterogeneity of urban area. For example, at 22.30°N, 114.17°E where the Kowloon Park locates, the LCZ map via VHR imagery successfully identifies the region as four small LCZ-13 scenes (bush, scrub), roughly an area of 306×306 m^2. But in a map via medium-resolution Sentinel-2 imagery, the region becomes too small to be correctly classified using scene classification. As a result, the map obtained via medium-resolution data is smooth but lacks details of urban context. The existence of green spaces like Kowloon Park has been proven to successfully mitigate the urban heat island effect [14]. Such details are essential in generating urban climate maps. Therefore, the usage of VHR imagery is crucial for LCZ mapping and is worth exploring.

Another noteworthy phenomenon is the capture of large roads and highways as LCZ-15 (bare rock or paved), e.g., the Stonecutters Bridge at 22.325°N, 114.12°E. These small features are often overlooked in LCZ mapping with medium-resolution imagery. But their high albedo nature leads to more solar radiation and is one of the causes of urban heat island [15]. Future investigations should look into whether the difference in handling paved materials leads to significant variations in weather research and forecasting models.

3.2. Multi-label LCZ Mapping Results

The multi-label classification results are shown in Table 2. For the three networks trained on the same domain (Hong Kong), the best result is obtained via the standard solution. It achieves an F_1 of 0.8550 and is ranked first by acc. The dilated solution is marginally after the standard solution and achieves a slightly smaller rankloss. The light-weight solution, though computation-efficient, has the worst performance among the three in terms of rankloss, acc, and F_1. But it has the best performance if evaluated by hloss. The hloss is more forgiving and only penalizes the individual labels, while other metrics evaluate the classification performance based on samples. The low hloss of the light-weight network means it assigns more incorrect labels to a given sample though one or several of the labels are correct. we show some multi-label classification examples in Fig. 4. The upper two are correctly classified. The label of lower left is compact high-rise and bush (LCZ-1/13), and the model predicts it as compact high-rise, bush, and rock/paved (LCZ-1/13/15). The label of lower

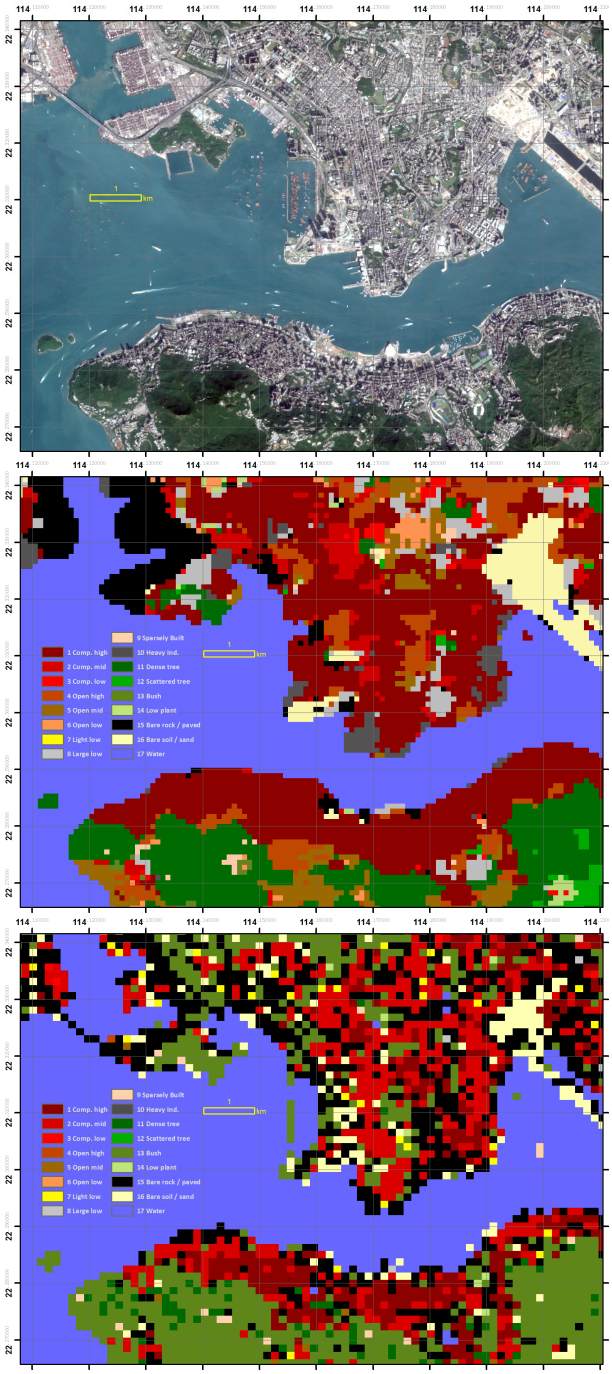


Fig. 3: Downtown single-label LCZ Maps and the reference image. Top to down: reference satellite image; LCZ map via 10 m imagery; LCZ map via VHR imagery.

Table 2: Classification. All: train on all samples including Hong Kong; Transfer: train on all other samples excluding Hong Kong. ↓: the smaller the better; ↑: the larger the better

	Train	Test	hloss ↓	rankloss ↓	acc ↑	F_1 ↑
Standard	Hong Kong	Hong Kong	.0418	.1421	**.7746**	**.8550**
	All	Hong Kong	.0510	.1899	.7085	.8022
	Transfer	Hong Kong	.0724	.2833	.5675	.6840
Light-weight	Hong Kong	Hong Kong	**.0399**	.1845	.7596	.8325
	All	Hong Kong	.0474	.1908	.7295	.8105
	Transfer	Hong Kong	.0687	.3047	.5864	.6833
Dilated	Hong Kong	Hong Kong	.0466	**.1313**	.7657	.8515
	All	Hong Kong	.0521	.1897	.7066	.7996
	Transfer	Hong Kong	.0787	.2828	.5394	.6662

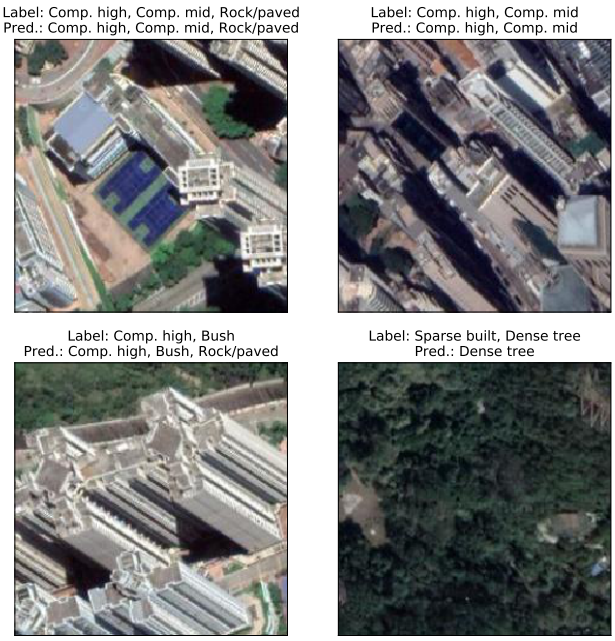


Fig. 4: Multilabel LCZs.

right is sparsely built and dense tree (LCZ-9/11), while the model neglects the sparsely built component.

The networks trained on all samples and the transfer learning still suffer from domain shift, although we expected the domain shift problem would be minimized since spectral response does not matter in three-channel VHR images. Other studies also found that a local solution is better for LCZ mapping in individual cities [2, 16]. Further investigations should look into the domain shift problem in LCZ mapping.

4. CONCLUSIONS AND FUTURE DIRECTIONS

We present a case study of multi-label local climate zone (LCZ) mapping in Hong Kong using scene classification techniques with very high resolution (VHR) imagery. By using VHR imagery, we can capture details of the urban landscape, e.g., small city parks, which can mitigate urban heat island effects and have an non-negligible impact on the local climate. These VHR images at a resolution better than 1 m provide more urban details than medium-resolution images with 10 m resolution. The hidden height information in the shading of buildings will benefit the detailing of LCZ mapping. The presented cases in Hong Kong already show promising advantages in using the VHR imagery.

With the multi-label techniques, we can classify an urban area into a combination of multiple climate zones, leading to subclass LCZ mapping. But the exponentially-increased output space requires models to handle the coexistence of multiple labels (climate zones). Future investigation should explore their relationships and analyze its performance in the generating LCZ maps to understand our local climate.

5. REFERENCES

[1] Yoo et al., "Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas using landsat images," *ISPRS J. Photogramm. Remote Sens.*, vol. 157, pp. 155–170, 2019.

[2] Liu and Shi, "Local climate zone mapping as remote sensing scene classification using deep learning: A case study of metropolitan china," *ISPRS J. Photogramm. Remote Sens.*, vol. 164, pp. 229–242, 2020.

[3] Bechtel et al., "Generating wudapt level 0 data–current status of production and evaluation," *Urban Clim.*, vol. 27, pp. 24–45, 2019.

[4] Gawlikowski et al., "On the fusion strategies of sentinel-1 and sentinel-2 data for local climate zone classification," in *IGARSS 2020*. IEEE, 2020.

[5] Le Saux et al., "2019 ieee grss data fusion contest: Large-scale semantic 3d reconstruction [technical committees]," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 33–36, 2019.

[6] Kotharkar and Bagade, "Local climate zone classification for indian cities: A case study of nagpur," *Urban Clim.*, vol. 24, pp. 369–392, 2018.

[7] Schneider and Woodcock, "Compact, dispersed, fragmented, extensive? a comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information," *Urban Stud.*, vol. 45, no. 3, pp. 659–692, 2008.

[8] Zheng et al., "Gis-based mapping of local climate zone in the high-density city of hong kong," *Urban Clim.*, vol. 24, pp. 419–448, 2018.

[9] Stefanakisl, "Web mercator and raster tile maps: two cornerstones of online map service providers," *Geomatica*, vol. 71, no. 2, pp. 100–109, 2017.

[10] Northcutt et al., "Confident learning: Estimating uncertainty in dataset labels," *arXiv preprint arXiv:1911.00068*, 2019.

[11] He et al., "Deep residual learning for image recognition," in *CVPR*, 2016, pp. 770–778.

[12] Wu and Zhou, "A unified view of multi-label performance measures," in *ICML2017*. PMLR, 2017, pp. 3780–3788.

[13] Zeiler, "Adadelta: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.

[14] Yang et al., "Assessing the impact of urban geometry on surface urban heat island using complete and nadir temperatures," *Int. J. Climatol.*, 2020.

[15] Mohajerani et al., "The urban heat island effect, its causes, and mitigation, with reference to the thermal properties of asphalt concrete," *J. Environ. Manage.*, vol. 197, pp. 522–538, 2017.

[16] Kim et al., "Developing high quality training samples for deep learning based local climate classification in korea," *NeurIPS2020 Workshop: AI for Earth Sciences*, 2020.