V-RSIR: A WEB-BASED TOOL AND BENCHMARK DATASET FOR REMOTE SENSING IMAGE RETRIEVAL

Dongyang Hou[1, 2], Huaqiao Xing[3]

[1] School of Geosciences and Info-Physics, Central South University, Changsha, Hunan, China-houdongyang1986@cumt.edu.cn
[2] College of Geography and Environment, Shandong Normal University, Jinan, Shandong, China
[3] School of Surveying and Geo-informatics, Shandong Jianzhu University, Jinan, Shandong, China- xinghuaqiao@126.com

KEY WORDS: Remote sensing image retrieval, Benchmark dataset, Volunteer, Web-based tool

ABSTRACT:

Benchmark datasets play an important role in evaluating remote sensing image retrieval methods. Current benchmark datasets are mostly collected through the Google Map API or other desktop tools. However, the Google Map API requires the users to have programming skills and other collection tools are not publicly available, which may hinder the development of new benchmark datasets. This paper develops an open access web-based tool V-RSIR to help users generating new benchmark datasets with volunteers for remote sensing image retrieval. Using this tool, a new benchmark dataset V-RSIR that contains 38 classes with at least 1500 images per class is created by 32 volunteers. A handcrafted low-level feature method and a deep learning high-level feature method are used to test the dataset. The evaluation results are consistent with our perception. This shows that the tool can help users effectively creating benchmark datasets for RSIR.

1. INTRODUCTION

The recent advances in satellite technology lead to a dramatic growth of remote sensing (RS) images (Chaudhuri et al., 2016; Tang et al., 2018). This has provided new opportunities for various RS applications; however, it also results in the significant challenge of retrieving RS images from a considerable volume of RS images (Ye et al., 2018;Shao et al., 2018). Therefore, developing remote sensing image retrieval (RSIR) approaches becomes one of the active and emerging research topics (Zhou et al., 2018).

In general, developing such RSIR approach requires extensive benchmark datasets to evaluate its performance (Zhou et al., 2018). The benchmark datasets need to include a huge amount of RS images labelled with categories (i.e., airplane, forest and freeway). At present, several RSIR benchmark datasets have been published on the web. For example, Zhou et al. (2018) presents a novel large-scale RS dataset named 'PatternNet' including 38 classes with 800 images per class; Shao et al. (2018) presents a multi-labelled dataset termed DLRSD including 17 classes with 2100 images. Besides, some benchmark datasets for RS image classification and object detection have also been used to evaluate RSIR approach in the past few years (Aptoula, 2014; Li et al., 2017; Zhou et al., 2017). For instance, the University of California, Merced dataset (UCMD) that is created initially for a land use/land cover classification (Yang and Newsam, 2010) becomes a most common benchmark dataset for RSIR.

Dataset	Images per class	Classes	Images	Resolutions(m)	Size	Format	References
PatternNet	800	38	30400	0.062–4.693	256*256	jpg	Zhou et al., 2018
DLRSD	-	21	2100	0.3	256*256	png	Shao et al., 2018
UCMD	100	21	2100	0.3	256*256	tif	Yang and Newsam, 2010
WHU-RS19	50	19	950	up to 0.5	600*600	jpg	Sheng et al., 2012
NWPU-RESISC45	700	45	31500	0.2-30	256*256	jpg	Cheng et al.,2017
RSSCN7	400	7	2800	-	400*400	jpg	Zou et al., 2015
SIRI-WHU	200	12	2400	2	200*200	tif	Zhao et al., 2016
AID	220-420	30	10000	0.5-8	600*600	jpg	Xia et al., 2017

Table 1. Publicly available datasets applied in RSIR

Table 1 lists several publicly available datasets which have been applied in RSIR. The maximum numbers for each class and all classes are 800 and 31500, respectively in the above datasets. These datasets are still small-scale datasets compared to ImageNet, which contains more than 14 million images and covers 20,000 categories (Russakovsky et al., 2015). On the other hand, the above RS datasets are mostly collected through the Google Map API or other desktop tools. However, the Google Map API requires the users to have programming skills and other collection tools are not publicly available. This may hinder the development of new large-scale benchmark datasets. Therefore, these have motivated increasing research interest on designing and developing an open access web-based tool for generating large-scale RSIR benchmark datasets.

This paper establishes an open access web-based tool termed V-RSIR. Concretely, this paper provides an overview of the V-

RSIR tool to date, which can help users generating new benchmark datasets with volunteers for RSIR. This includes image single-label, image cropping, image editing, image review, image statistics, spatial distribution of images and image sharing. Besides, Using this tool, a new benchmark dataset which is also termed V-RSIR is created by 32 volunteers. The new benchmark dataset contains 38 classes with at least 1500 images per class.

A longer term goal of V-RSIR tool is to simultaneously offer image single-label and multi-label function, and build a really large-scale benchmark dataset for RSIR as good as ImageNet.

The remainder of this paper is organized as follows. Section 2 overviews the V-RSIR tool and Section 3 describes the new V-RSIR benchmark dataset, followed by conclusions.

2. THE V-RSIR TOOL

2.1 V-RSIR Architecture

The V-RSIR tool is a web application that annotates and crops RS images from online images maps by worldwide volunteers for constructing RSIR benchmark datasets. Currently, the V-RSIR tool is only available in English. It is now hosted in the Aliyun server and is available at http://www.geoinfobar.com:5321.
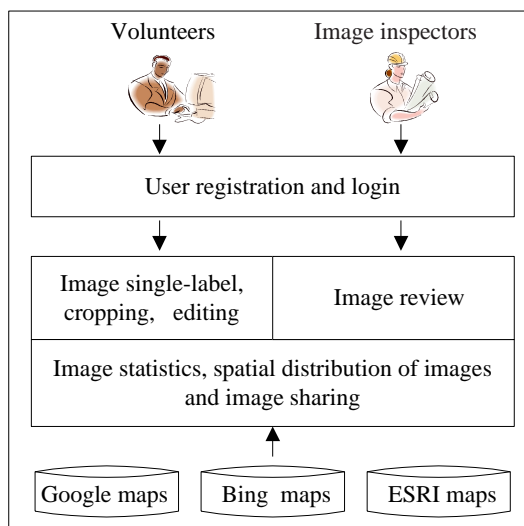

Figure 1. Conceptual architecture of the V-RSIR tool

Figure 1 presents the conceptual architecture of the V-RSIR tool. From bottom to top, the V-RSIR tool is divided into three tiers. Tier 1 is the data layer, in which Google maps, Bing maps and ESRI (Environmental Systems Research Institute) maps and other online maps are served as RS image sources. These online maps are directly integrated into the V-RSIR tool through their service addresses. The top tier is the user group including volunteers and image inspectors. The volunteers can label, crop and edit RS images. Besides, they also can browse the quantity statistics and spatial distribution of cropped images and download the inspected images. In addition to the above permissions, the image inspectors are mainly responsible for checking whether the category is labelled correctly by volunteers. The middle tier is the business logic of the V-RSIR tool, including user registration, login, image single-label, cropping, editing, review, quantity statistics, spatial distribution, sharing and so on.

2.2 Main functions of the V-RSIR tool

The V-RSIR tool is implemented based on an open source library of OpenLayers. This section outlines the main functions in Figure 1.

2.2.1 Image single-label and cropping: When the user successfully logs, the tool will skip to the main pages of image single-label and cropping (Figure 2).
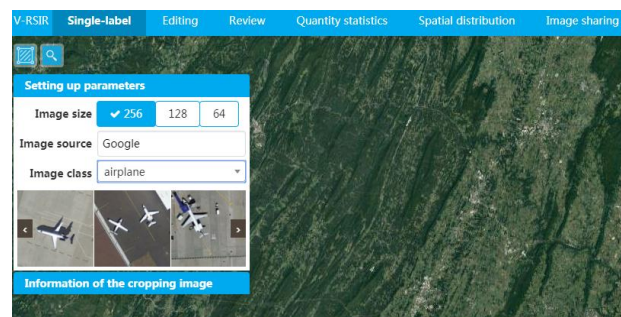

Figure 2. Interface of image single-label and cropping

Users should first set image size, image source, image category in the left setting panel. Then, users can enter category keywords and place names into the search box to retrieve and locate similar scenes. This is implemented based on Nominatim (https://nominatim.openstreetmap.org). For example, if users want to label baseball field category, they can select the baseball field category in the left setting panel and enter the keywords of " baseball Washington" into the search box, some similar scenes will be displayed in the drop-down box (Figure 3). When users determine the correct category in the located scenes, they can click the square icon in the top left corner of the page to draw a yellow bounding box around the correct scene (Figure 4). If the scene has been labelled by other users, a hint panel with thumbnails will pop up in the bottom right corner of the page (Figure 4) and the details of the labelled image will appear if users click the thumbnail (Figure 5). If the scene has not been labelled, users can click the 'submit' button (Figure 4) to crop it and store it into the database.
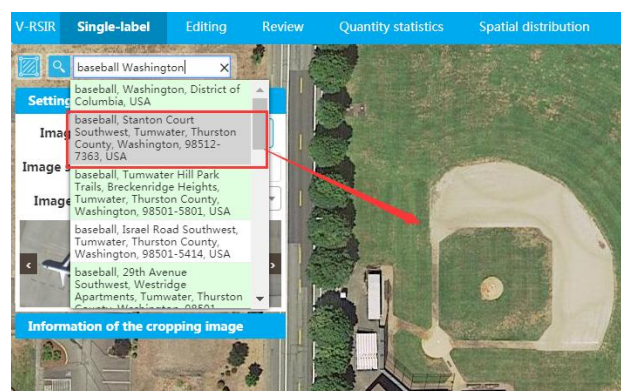

Figure 3. Interface of search example

Figure 4. Interface of cropping images

Figure 5. Interface of details of labelled images

2.2.2   Image editing: The volunteers can manage labelled images by themselves in the page of image editing (Figure 6). Once they find images are labelled incorrectly, they can delete them individually or in batches through 'delete' button or 'bulk delete' button.
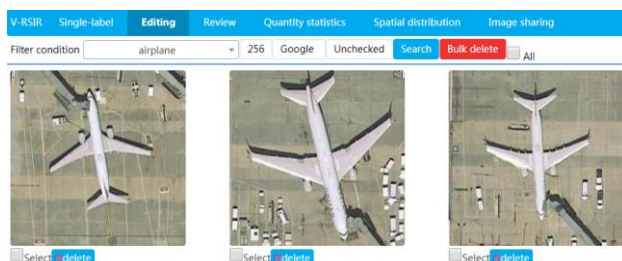

Figure 6. Interface of image editing

2.2.3   Image review: Occasionally, volunteers may make mistakes due to their negligence and not all volunteers follow the instructions to label images. The solution to these issues is have professionals to review the images again. Therefore, the V-RSIR tool sets some image inspectors to review the images. The image inspectors can delete incorrect images individually or in batches in the page of image review (Figure 7). Besides, the correct images should be submitted to the benchmark dataset by clicking 'Check' button (Figure 7).
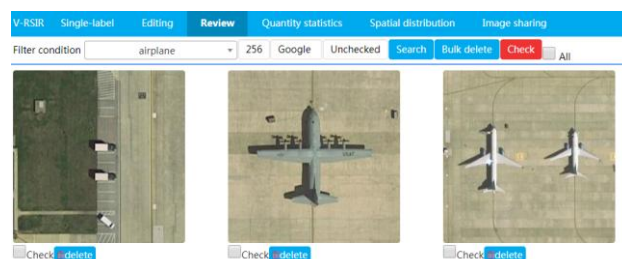

Figure 7. Interface of image review

2.2.4   Image statistics: Users can browse the number of checked and unchecked images through the 'Statistics' button (Figure 8). The result can be displayed in the form of histogram or line chart by using the library of ECharts (https://echarts.baidu.com/) .
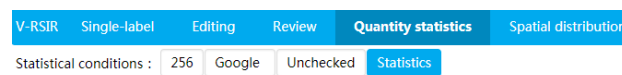
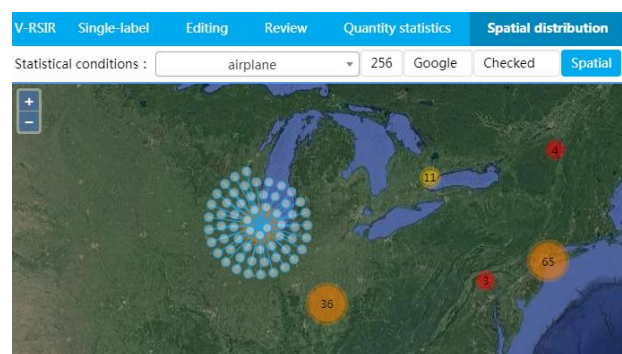

Figure 8. Interface of image statistics

Figure 9. Interface of spatial distribution

2.2.5   Spatial distribution of images: Users can browse the spatial distribution of checked and unchecked images through the 'Spatial' button (Figure 9). The result is displayed in the form of clusters. After clicking one circle of the clusters(Figure 9), the map will locate to the position of the corresponding image.

2.2.6   Image sharing: Image sharing provides sharing of the checked datasets via the web. At present, the labelled data can be acquired by download link of Baidu cloud disk

3.   WALK-THROUGH EXAMPLE: V-RSIR BENCHMARK DATASET

In this section, a walk-through example of constructing a new benchmark dataset termed V-RSIR is presented to demonstrate the usefulness of the V-RSIR tool.

The walk-through example adopts the classification scheme of PatternNet dataset, which covers 38 classes, such as airplane,

baseball field, basketball court and so on (Zhou et al., 2018). Beside, the example requires the labelled images with the size of 256*256 pixels from Google maps. With the advantage of working in a university, 32 volunteers and 6 image inspectors participate with remote sensing background spend 2 weeks with labelling 30 minutes to 1 hours per day for the new V-RSIR benchmark dataset. In this example, each volunteer is assigned with 1 or 2 categories annotation work and each image inspector is assigned with 6 or 7 categories of inspection work. Then, volunteers use the function of image single-label to collect images and inspectors uses the function of image review to check images labelled by volunteers

Currently, the new V-RSIR benchmark dataset covers 38 classes and contains 59404 images with at least 1500 per class. The spatial resolutions of these images range from 0.075 to 9.555 meters. The dataset takes up 8.70 GB of storage. Figure 10 represents some example images from the V-RSIR benchmark dataset.

(a) airplane   (b) baseball field

(c)basketball court   (d) beach

(e) bridge   (f) crosswalk
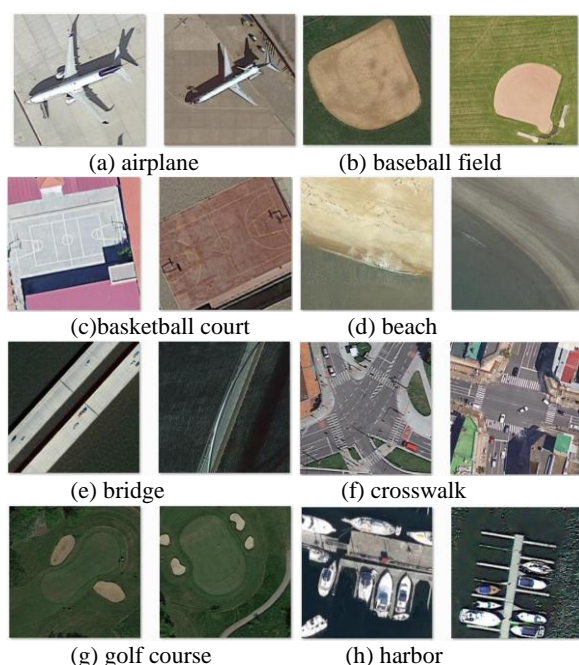
(g) golf course   (h) harbor
Figure 10. some example images from the V-RSIR dataset

Figure 11 represents the spatial distribution of the V-RSIR dataset. The images in this dataset are mainly distributed in North America, Europe and South America.

Figure 11. Spatial distribution of the V-RSIR dataset

A handcrafted low-level feature method of HSV color (hue, saturation, value) and a deep learning high-level feature method of DenseNet (dense convolutional network) are selected to verify its ability of evaluating different RSIR methods. We randomly select 50 images from each class of the VGoogle-RSIR dataset as the test set and the remaining images are used as the training set. In other words, there are 1900 images in the test set and 57504 images in the training set.

Average normalized modified retrieval rank (ANMRR), mean average precision (mAP) and precision at k (Pk where k is the number of retrieved images) are used to evaluate the retrieval performance. For ANMRR, lower values mean better performance and higher values mean better performance for the other two metrics (Zhou et al., 2018). The averages of all queries about the three metric are shown in Table 2.

Methods	ANMRR	mAP	p5	p10	p20	p50	p100	p1000
HSV	0.80	0.13	0.59	0.54	0.48	0.40	0.34	0.13
DenseNet	0.29	0.63	0.94	0.93	0.92	0.90	0.88	0.71

Table 2. The results of the two methods on the VGoogle-RSIR dataset.

The performance of HSV is worse than the DenseNet method DenseNet, which is consistent with our perception. This illustrates that the dataset labelled by the V-RSIR tool can effectively evaluate the retrieval performance.

4. CONCLUSIONS

This paper presents an open access web-based tool V-RSIR and it is intended to help users generating new benchmark datasets for RSIR. The tool offers functions of image single-label, cropping, editing, review, quantity statistics, spatial distribution, sharing and so on. To evaluate the usability of the tool, 32 volunteers and 6 image inspectors are organized to label the images by using the tool. Ultimately, a new benchmark dataset V-RSIR covering 38 classes with least 1500 images per class is constructed. This demonstrate the effectiveness and applicability of the V-RSIR tool. Future work will concentrate on improving the functions of image sharing and image multi-label.

ACKNOWLEDGEMENTS

REFERENCES

Chaudhuri, B., Demir, B., Bruzzone, L., & Chaudhuri, S., 2016. Region-based retrieval of remote sensing images using an unsupervised graph-theoretic approach. *IEEE Geosci. Remote Sens. Lett.*, 13(7), 987-991.

Tang, X., Zhang, X., Liu, F., Jiao, L., 2018. Unsupervised Deep Feature Learning for Remote Sensing Image Retrieval. *Remote Sens. , 10*, 1243.

Ye, F., Xiao, H., Zhao, X., Dong, M., Luo, W., Min, W., 2018. Remote Sensing Image Retrieval Using Convolutional Neural Network Features and Weighted Distance. *IEEE Geosci. Remote Sens. Lett.*, (99), 1-5.

Shao, Z., Yang, K., Zhou, W., 2018. Performance Evaluation of Single-Label and Multi-Label Remote Sensing Image Retrieval Using a Dense Labeling Dataset. *Remote Sens. 10*, 964.

Zhou, W. , Newsam, S. , Li, C. , Shao, Z., 2018. Patternnet: a benchmark dataset for performance evaluation of remote sensing image retrieval. I*SPRS J. Photogram. Remote Sens.*, 145(A), 197-209.

Li, Y. , Zhang, Y. , Huang, X. , Zhu, H. , Ma, J., 2017. Large-scale remote sensing image retrieval by deep hashing neural networks. *IEEE Trans. Geosci. Remote Sens.*, (99), 1-16.

Aptoula, E., 2014. Remote sensing image retrieval with global morphological texture descriptors. *IEEE Trans. Geosci. Remote Sens.*, 52(5), 3023-3034.

Yang, Y., Newsam, S., 2010. Bag-of-visual-words and spatial extensions for land-use classification. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. San Jose, California, USA. doi. org/10.1145/1869790.1869829

Sheng, G., Yang, W., Xu, T., Sun, H., 2012. High-resolution satellite scene classification using a sparse coding based multiple feature combination. *Int. J. Remote Sens.*, 33(8), 2395-2412.

Cheng, G., Han, J., Lu, X., 2017. Remote sensing image scene classification: benchmark and state of the art. *Proc. IEEE*, 105(10),1865-1883.

Zou, Q., Ni, L., Zhang, T., Wang, Q., 2015. Deep Learning Based Feature Selection for Remote Sensing Scene Classification. *IEEE Geosci. Remote Sens. Lett.*, 12(11), 2321-2325.

Zhao, B., Zhong, Y., Xia, G. S., Zhang, L., 2016. Dirichlet-derived multiple topic scene classification model for high spatial resolution remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.*,54(4), 2108-2123.

Xia, G. S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., ... & Lu, X. (2017). AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Trans. on Geosci. Remote Sens.*, 55(7), 3965-3981.

Zhou, W., Newsam, S. , Li, C., Shao, Z., 2017. Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval. *Remote Sens.*, 9(5), 489.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al., 2015. Imagenet large scale visual recognition challenge. *Int. J. Computer Vision*, 115(3), 211-252.