An Improved Framework for Airport Detection
Under the Complex and Wide Background

Ning Li, Liang Cheng, Chen Ji, Shengkun Dongye, and Manchun Li

Abstract—As an important transportation facility, airports substantially affect the economic lives of people. However, the full extraction of airports located in a vast area is concerning. The size of an airport in a previous wide area detection framework is relatively large and has a strong saliency in remote sensing images, whereas the contradiction between a complex geographical background and a small airport size has yet to be resolved. In this study, we propose a set of automatic detection frameworks to realize efficient detection for various airports in nine Indian states/union territories under the condition that only runway samples are labeled. Preliminary extraction of runway features is performed with a high F1 and recall rate, and teacher nodes judge and guide the results. Next, the output is connected to classification and segmentation for outlier elimination and pixel extraction to locate the runways. For the study area, the proposed framework airport retention rate (RR) is 92.7%, with the false alarm reduction rate (FARR) reduced by a maximum of 95.3%. A total of 192 airports are discovered, and the effective airport growth rate (GR) is 47.4%. Compared to previous work, RR, GR, FARR, and run efficiency increased by 2.2%, 16.0%, 4.5%, and 432.5%, respectively, with more small- and medium-sized airports detected. Furthermore, the framework is tested in Japan, and 155 airports are detected. Thus, the proposed framework effectively improves detection capability for small- and medium-sized airports in large-scale areas and updates the airport database.

Index Terms—Airport detection, complex and wide area, deep learning, remote sensing, small airport.

I. INTRODUCTION

As important aviation infrastructure, airports play an indispensable role in civil and military aspects, including transportation, economy, supply, communication, and others [1]. More comprehensive traffic planning scenarios and analysis of human mobility patterns can be derived from studies of global airport distribution [2], [3]. The quantity and quality of existing public airport databases are imperfect, mainly reflecting incomplete data and deviations in spatial location. Incomplete data indicate inadequate coverage of all airports in the region, whereas spatial information deviation indicates the inaccurate spatial location of an airport in the database. The development of technologies such as high-resolution remote sensing images and artificial intelligence has provided a technical basis for accurately mining airport locations in vast areas. Considering the complexity of remote sensing images in a vast area, target mining using a single model is challenging. Along with deep-learning techniques, a proper geographical analysis can substantially reduce the misjudgment rate of the framework in the process.

Applying deep-learning technology can considerably improve the efficiency and accuracy of object detection in vast-area remote sensing images [2], [3], [4], [5], [6], [7], [8]. Li et al. [4] used YOLOv3 [9] to identify unknown airport locations effectively from high-resolution images in China’s Yangtze River Delta (YRD). However, the size of the airports in the study area [4] is relatively large, which reduces the difficulty of detection; thus, the method is not applicable in areas with more complex backgrounds and smaller airports. Zeng et al. [2] first extracted airport candidate areas using FROM-GLC10 production [10] and used a well-trained model of faster-RCNN [11] to determine aircraft location. However, it is seriously affected by the accuracy of auxiliary data. Then, with Sentinel-2 imagery, they extracted the built-up land area of Chinese airports by deep learning [3]. Marcum et al. [7] detected missile positions based on RESNET-101[12] and spatial analysis in China Research Area; the method was 81 times faster than visual search. Yuan et al. [6] mapped solar panels based on deep convolutional networks in imagery covering 200 km² in two cities. Yu et al. [5] developed a framework (DeepSolar) to analyze satellite imagery, and a comprehensive solar database was constructed for the contiguous US. Although the development of deep-learning technology has improved the efficiency of object detection in vast regions, extraction results with huge differences were affected by regional scale, feature analysis methods, and object type. Therefore, the spatial information of airports in a certain region with complex and wide features, such as the Indian region, must be located quickly and efficiently.

To apply deep learning in the full extraction of airport information in open areas, different methods must be integrated to detect, classify, and segment unforeseen images. Most studies on airport detection are based on single images [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], and only a few studies have...
Fig. 1. Study area.

| TABLE I |
|———|———|———|———|
| **Type** | **Source** | **Number** | **Usage** |
|———|———|———|———|
| Airports | OA | 137 | Teacher- node |
| Airports | Global | 2043 | Training |
| Airports | [34] | 940 | Testing |
| Remote sensing images tiles | Google Maps | 25481088 | Basic images |
| Roads/water/railways | OSM | - | Teacher- node |
been conducted for vast areas [2], [3], [4], [34] and for detecting or validating the use of complicated processing procedures or high algorithm requirements. Classic object detection algorithm frameworks, such as faster-RCNN [11] and YOLOv3 [9], have different advantages in terms of precision and speed. In addition, researchers [35], [36], [37] have introduced a transformer of the self-attention mechanism in computer vision tasks to achieve a high-efficiency process. For classifying one-class targets, some researchers regarded them as positive and negative two-class classification problems [4], [5], [7], [8], whereas others performed anomaly detection processing by constructing a one-class classifier [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], which only learns through one class of samples. The rapid development of the semantic segmentation model [49], [50], [51], [52], [53], [54], [55] also provides a deep-learning method for fine extraction of airport objects.

Geographical feature analysis can effectively improve the efficiency and reliability of remote-sensing airport detection. Many publicly available geographical data products provide useful prior knowledge for airport mining in vast areas, such as global land-cover products [10], [56], [57], Open Street Map

**TABLE II**

| Models   | True | False | Precision | Recall | F1   |
|----------|------|-------|-----------|--------|------|
| Faster-RCNN | 640  | 10    | 98.5%     | 68.1%  | 0.805|
| YOLOv3   | 795  | 6     | **99.3%** | 84.6%  | 0.914|
| YOLOv5   | 671  | 20    | 97.1%     | 71.4%  | 0.823|
| CenterNet| 594  | 20    | 96.7%     | 63.2%  | 0.764|
| DETR     | 821  | 26    | 96.9%     | **87.3%** | **0.918** |

Fig. 2. Improved framework in the study area.

Fig. 3. Calculation process of the integrated detector.
TABLE III
EFFICIENCY STATISTICS

| Methods                              | Steps            | RR (%) | FARR (%) | GR (%) | Time (s) |
|--------------------------------------|------------------|--------|----------|--------|----------|
| Runway extraction                    |                  | 99.3   | -        | -      | 55 187   |
| Elimination of airport outliers      |                  | 93.4   | 89.1     | -      | 20 926   |
| Runway segmentation                  |                  | 92.7   | 95.3     | -      | 4672     |
| Validation                           |                  | 92.7   | 95.3     | 47.4   | 2145     |
| Total                                |                  | 100    | 100      | 47.4   | 82 930   |

Traditional visual research          -        | 100    | 47.4    | 7 015 693 |

Fig. 4. Change of FARR.

Fig. 5. Examples of runway segmentation. (a) Amreli Airport. (b) Salem Airport. (c) Gwalior Airport. (d) Kishangarh Airport.

TABLE IV
COMPARISON RESULTS

| Frameworks | RR (%) | GR (%) | FARR (%) | Run-time (s) |
|------------|--------|--------|----------|--------------|
| [2]        | 32.8   | 2.2    | 8.3      | -            |
| [4]        | 90.5   | 31.4   | 90.8     | 358 694      |
| Ours       | 92.7   | 47.4   | 95.3     | 82 930       |

(OSM) [58], and human settlement-related products [59]. Zeng et al. explored airports with land-cover products with impervious surface properties [2]. However, this method is only effective for larger airports, and smaller airports may be missed owing to issues, such as product resolution and accuracy. Jing et al. [60] comprehensively considered the OSM Road, OSM Farmland, OSM bridge, SRTM1, AW3D30, From-GLC10, and other data to make constraints based on empirical knowledge, effectively improving the detection of dams in Japan. Li et al. [4] performed spatial analysis on detection boxes and road network data in the OSM to narrow the candidate area. Therefore, as far as airports are concerned, detection that considers reasonable experience and knowledge in a vast area can effectively improve the problems of low efficiency and low precision. However, the detection of small airports in complex areas still needs improvement.

In this study, we propose a set of airport mining frameworks that integrate different deep learning and geographical analysis algorithms to realize the extraction of various types (large, medium, and small) of airports in vast and complex areas. In view of the inability of previous methods to meet more complex geographical backgrounds and smaller airport types, the main contributions of the proposed framework are as follows.

1) Under the premise that small- and medium-sized airports cannot be effectively detected in vast areas, an airport extraction framework was improved, breaking through the limitations of a single solution. Our framework was tested in Japan, showing that it is effective and fast.

2) Datasets of airport locations (airports with runway features) in two regions of India and Japan were updated.

The remainder of this article is as follows: Section II introduces the study area and materials and describes the framework in detail. Results are presented in Section III, and a discussion is given in Section IV. Section V concludes this article.

II. MATERIALS AND FRAMEWORK

A. Study Area and Materials

Airports, as one of the most important means of transportation, have a strong correlation with the economy and population. We selected nine Indian states/union territories that rank high in the regional economy and population density as study areas, and
Fig. 6. Distribution of the OA, newly found, and suspected airports.

their distribution is shown in Fig. 1. The total area is 1,881,488 km².

From the open database, OurAirports (OA; www.ourairports.com), 149 Indian airport locations, including large, medium, small, and closed airports, were obtained. After visual inspection, 137 airports with evident runway characteristics were regarded as positive teacher nodes for validation. To obtain sample data for airport detection, we randomly labeled 2043 airports worldwide for training. A total of 940 airports with confidence levels > 99% in the results of a previous study [34] were extracted for testing. In addition, Google images tiles (Level 17) of the study area were acquired as basic image data (training, testing, and detection). OSM (http://download.geofabrik.de/) and other databases were applied for other teacher nodes (see Table I). The experiment was based on the RTX 2080Ti GPU cards, and the program was realized based on Kubernetes.

B. Framework

The method proposed in this study to detect airports located in nine Indian states/union territories is shown in Fig. 2.

The runway of an airport is the most evident feature in remote sensing images and was taken as the main target of airport detection. We selected 2983 global airport samples, of which 2043 were used for training and 940 for testing. The number of training samples was expanded to 16,994 by rotation and...
clipping. The first key to airport mining is to ensure airport recall while ensuring precision; therefore, many deep-learning algorithm frameworks, such as faster-RCNN [11], YOLOv3 [9], YOLOv5 [63], detection transformer (DETR) [36], and CenterNet [61], are applied for comparative experiments. By synthetically comparing precision, recall, and F1, the model effect was analyzed to obtain the optimal model or model combination.

In addition, numerous airport outliers remained because of the study area’s vast and extremely complex surface. The misidentification of similar geographical objects is the main reason for the high false alarm rate. To analyze the image with geospatial data and instruct the classification model on whether the features were positive or negative, the roads, water polylines/polygons, and railways in existing databases (such as OSM) were regarded as teacher nodes for negative sample generation. We filtered outliers by extracting some correct and incorrect samples as supervised data. Based on global samples, part of the verified airports in OA were selected as positive samples, and some were extracted from teacher nodes for negative samples. The specific steps are as follows:

**Step 1:** Spatialization First, we convert the result box identified by the DETR model from image space to geographic space; precisely, the pixel coordinates of the box are converted to geospatial coordinates

$$B_{\text{lat0,lng0,lat1,lng1}} = f_{\text{convert}}(B_{\text{pl,pt,pr,pb}}, z)$$

where $B_{\text{pl,pt,pr,pb}}$ is the pixel coordinates (left, top, right, bottom) of the box, and $z$ is the image level, 17 in this study. $B_{\text{lat0,lng0,lat1,lng1}}$ are the geographical coordinates collection of the box. $f_{\text{convert}}(\cdot)$ is the conversion process (refer to https://developers.google.com/maps/documentation/javascript/coordinates).

**Step 2:** Match The spatialized geographic boxes are spatially matched with the teacher node information. The boxes matching OA nodes are regarded as positive, and the boxes matching other teacher nodes are regarded as negative

$$\begin{align*}
   & \begin{cases}b \in \text{positive}, b \cap \text{TeacherNode}_{\text{OA}} \neq \text{NULL} \\
   & \begin{cases}b \in \text{negative}, b \cap \text{TeacherNode}_{\text{other}} \neq \text{NULL}
   \end{cases}
   \end{cases}
\end{align*}$$

where $b$ is each detection box.

**Step 3:** Sample Generation The geographic boxes of the two classes are inverted into pixel coordinates boxes, and their positions in the image space are calculated to generate a collection of image samples. $f_{\text{invert}}(\cdot)$ is the inversion process.

$$B_{\text{pl,pt,pr,pb}} = f_{\text{invert}}(B_{\text{lat0,lng0,lat1,lng1}}, z).$$
The study area was divided into 853 mission areas, each containing 625 images to be detected, totaling 533 125 images. To improve the detection precision and recall rate in this area substantially, we constructed an extraction algorithm (2) based on the two detectors mentioned above.

\[
R = \sum_{i=1}^{m} Y_i \left( D_i \cap Y_i \neq \emptyset \right) + \sum_{j=1}^{n} D_j \left( Y_j = \emptyset, \text{score} > \alpha \right)
\]

(5)

where \( R \) is the result of runway detection, \( m \) is the number detected concurrently by the two, \( \alpha \) is the threshold of the DETR detection score, \( n \) is the number detected by DETR only, and the scores greater than \( \alpha \), \( D_i \), and \( Y_i \) are the extraction results of the DETR and YOLOv3 models on the \( i \)th image, respectively.

**B. Runway Detection**

Fig. 3 illustrates the calculation process of the integrated detector; Fig. 3(c) and (g) shows the extraction results of the airport effective area under the two conditions \( D_i \cap Y_i \neq \emptyset \) and \( Y_i = \emptyset, \text{score} > \alpha \), respectively. Each thread processed 43 task areas of 853 missions for a total of 20 threads. In (5), the value of \( \alpha \) affects the complexity and efficiency of subsequent task processing; simultaneously, to fully ensure the known airport recall rate, we calculate the \( \alpha \) threshold of each airport in the OA database. When \( \alpha \) is 0.98, all known airports are recalled, which this article recommends. In total, 83 522 candidate boxes were extracted from the DETR model, and 26 089 were extracted from the YOLOv3. \( R \) is 40 485 calculated by (2). DETR has a 100% retention rate (RR) of the airports in OA, whereas YOLOv3 and \( R \) have 97.8% and 99.3%, respectively. \( m \) and \( n \) are 15 385 and 22 939, respectively.

**C. Elimination of Airport Outliers**

Based on the teacher nodes of roads, water polygons/polylines, railways, and airports in OA, the detection results were analyzed. To fully explore the potential unknown airports, we calculated the RR of airports in the OA database under different \( \beta \) and determined the value of \( \beta \) to be 0.1 with max RR. The shortest airport length was approximately 500 m, and the narrowest airport width was approximately 30 m, corresponding to the values of \( \gamma \) and \( \delta \), respectively. The change in false alarm reduction rate (FARR) is shown in Fig. 4. After the above process, the overall airport RR remained at 99.3%, and the overall FARR was reduced by a maximum of 46.0%.

\[
\text{FARR} = \frac{\text{number}_b - \text{number}_r}{\text{number}_b} \times 100\%
\]

(6)

where \( \text{number}_b \) is the number of detection boxes, and \( \text{number}_r \) is the number of reserved.

Seventy-one airport samples and 3801 abnormal samples verified by teacher nodes were selected as supervised samples for anomaly detection. A total of 15 epochs were trained based on the ImageNet pretraining model. The loss, accuracy of training, and validation were 0.0669, 0.9788, 0.0972, and 0.9684, respectively. To improve the RR of known airports, we set the confidence score of the model to 0.1. After detection, \( m \) was
updated to 2625 and \( n \) to 1145, with an RR of 93.4\%, and the invalid candidate boxes were reduced by 89.1\%.

### D. Runway Segmentation

Through 50 trained epochs based on DDRNet-39[62] pre-training model of ImageNet on airport samples, the training loss and accuracy were 0.0015 and 0.9961, respectively. After the runway segmentation, \( m \) was updated to 1632, \( n \) was updated to 181, the RR was 92.7\%, and the number of invalid candidate areas was further reduced by 51.9\%. In addition, we fused the boxes with an overlapping degree in the recognition process, and finally, \( m \) and \( n \) were updated to 508 and 144, respectively. The example results of partial airport runway segmentation of correct and incomplete segmentations are shown in Fig. 5.

### E. Extraction Results

A total of 192 airports were detected in the study area, of which 127 were verified in the OA database and 65 were newly found airports. As for the distribution of newly found airports, 60 were in \( m \), and 11 were suspected; five were in \( n \), and two were suspected (see Fig. 6). The RR and effective growth rate (GR) of airports in this framework were 92.7\% and 47.4\%, respectively.

### F. Efficiency Statistics

We presented the RR, the FARR of invalid candidate areas, the effective GR, and elapsed time of each step of the framework. Our method is 85 times better than any traditional visual search. At the same time, it maintains a high RR and GR and realizes the deep extraction of airports in a large-scale area.

\[
RR = \frac{\text{airports}^{\prime}_{\text{OA}}}{\text{airports}_{\text{OA}}} \times 100\% \tag{7}
\]

\[
GR = \frac{\text{airports}_{d} - \text{airports}^{\prime}_{\text{OA}}}{\text{airports}_{\text{OA}}} \times 100\% \tag{8}
\]

where \( \text{airports}^{\prime}_{\text{OA}} \) is the number of detected airports in the OA database, \( \text{airports}_{\text{OA}} \) is the number of airports in the OA database, and \( \text{airports}_{d} \) is the number of detected airports from our research area.

The efficiency statistical results obtained according to formulas (6)–(8) are shown in Table III. The runtime of our framework is 82,930 seconds.

### IV. DISCUSSION

#### A. Test Results for This Framework

Fig. 7 shows the test result of the improved framework in Japan. We divided the Japanese region into 684 detection task areas and identified them using the proposed processing framework. Finally, we obtained a total of 155 airports with obvious runways with an 85.4\% precision rate, 137 of which were recorded in the OA database. The recall rate of the database was 95.6\%, and nine known airports were lost. Eighteen airports were newly found relative to OA (blue dots in Fig. 7), where the facilities in the airports are relatively single, generally with only runways and slightly small sizes in remote sensing images. The entire process took 3 hours and 11 minutes using the trained models on the Kubernetes platform. The runway extraction and segmentation models were also directly transferred to the Japanese region and achieved the above results. The result not only proves the efficacy of the airport detection framework but also that the trained models have a strong generalization ability.

The airports with runways in Japan are available at https://github.com/ChanganLeo/JapanAirports.

#### B. Comparisons of Different Frameworks

There are still relatively few studies on airport detection and extraction in very large areas. We compared our framework for airport detection with several meaningful works, including the frameworks of [2] and [4]. We applied their methods to the research area of this article and obtained the results in Table IV.

As can be seen from Table IV, our framework (bold entities) is significantly improved compared with the research framework in [4]; RR, GR, FARR, and run efficiency increased by 2.2\%,
16.0%, 4.5%, and 432.5%, respectively, with more small- and medium-sized airports detected. The framework in [2] uses a large number of auxiliary data as candidate areas, many processing processes rely on manual operation, and it takes aircraft as the characteristics of airports, so many airports are left out.

C. Candidate Areas
As different regions have different economic development models and airport construction levels, this study aimed to extract airports fully from the research area, including large, medium, small, and closed airports, as well as other airports with no prominent characteristics. To improve the efficiency of airport search in a vast area, some scholars searched for aircraft in the impervious surface area [2]; however, they often lost numerous targets because of the accuracy of the impervious surface data when applied to our study area. Fig. 8 shows several coverage examples of the FROM-GLC10 impervious surface products. Before, we took a scene classifier in conjunction with spatial analysis to obtain the candidate areas in the YRD region of China [4], but it is time-consuming. We calculated the average length and width of airports in Madhya Pradesh and YRD to analyze their airport saliency. The average length and width of the former are about 1650 and 37 m, respectively, whereas those of the latter are 2695 and 55 m, respectively. Therefore, after spatial analysis in [4], the recall rate in YRD can still maintain a high level, whereas the recall rate in MP is difficult to ensure, thus bringing additional noise to the semantic information of tiles scenes. Therefore, we adopted an integrated detection method based on deep learning to ensure that the airport is recalled fully.

D. Incorrect Extraction Analysis
Among the final extraction results in the study area, many misextractions of roads, water polylines/polygons, and railways were not removed; long and narrow buildings, farmland, and other features were also similar to runways. Concerning airport semantics, a boundary that entirely distinguishes the runway from unknown features in a vast unknown scanning area is difficult to obtain. From the geographical analysis perspective, several low-level roads, water bodies, and other features in the study area could not be effectively updated into geographic data, resulting in errors in teacher nodes. Therefore, the integrity and accuracy of geographical data are important factors that affect the extraction results of this framework. At the same time, the effectiveness of the method is demonstrated with high FARR.

E. Lost Airports
A total of ten airports were not successfully extracted (see Fig. 9), and the loss rate was 7.3%. Lost runways comprised most small airports with a single runway, and the material of some runways was degraded to the soil. The detection ability of the framework for some small and indistinguishable airports (short and narrow runways) warrants improvement.

V. Conclusion
We improved an airport extraction framework for a vast area, and the framework effectively mined 192 airport spatial locations in India. In addition, the framework was tested in Japan, and 155 airports were detected. Compared to previous work, RR, GR, FARR, and run efficiency increased by 2.2%, 16.0%, 4.5%, and 432.5%, respectively, with more small- and medium-sized airports detected. It effectively improved the detection capability of small and medium-sized airports in large-scale areas. The comprehensive use of deep learning and geographical analysis can continuously narrow the candidate area, thereby improving detection efficiency. Compared with manual work, our framework was 85 times more efficient. However, the extraction ability must be improved for some small and narrow airports.

Given the vast research area, more work needs to be done to get detection models with high recall and accuracy: at the same time, more reasonable geographic analysis methods should be developed. Another limitation is the small sample size, as this study was only focused on the runways of airports. For more complex application scenarios, the theoretical methods in the open world may have a substantial effect on this study. Establishing an object classification framework to construct more definite categorical boundaries can facilitate detection tasks, anomaly detection, and others.

REFERENCES
[1] S. Li, Y. Xu, M. Zhu, S. Ma, and H. Tang, “Remote sensing airport detection based on end-to-end deep transferable convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 10, pp. 1640–1644, Oct. 2019, doi: 10.1109/LGRS.2019.2904076.
[2] F. Zeng et al., “A hierarchical airport detection method using spatial analysis and deep learning,” Remote Sens., vol. 11, no. 19, Jan. 2019, Art. no. 2204, doi: 10.3390/rs11192204.
[3] F. Zeng, X. Wang, and M. Zha, “Extracting built-up land area of airports in China using Sentinel-2 imagery through deep learning,” Geocarto Int., pp. 1–21, 2021, doi: 10.1080/10106049.2021.1983034.
[4] N. Li et al., “Framework for unknown airport detection in broad areas supported by deep learning and geographic analysis,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 6326–6338, 2021, doi: 10.1109/JSTARS.2021.3089811.
[5] J. Yu, Z. Wang, A. Majumdar, and R. Rajagopal, “DeepSolar: A machine learning framework to efficiently construct a solar deployment database in the United States,” Joule, vol. 2, no. 12, pp. 2605–2617, Dec. 2018, doi: 10.1016/j.joule.2018.11.021.
[6] J. Yuan, H. H. L. Yang, O. A. Omitsienu, and B. L. Bhaduri, “Large-scale solar panel mapping from aerial images using deep convolutional networks,” in Proc. IEEE Int. Conf. Big Data, 2016, pp. 2703–2708, doi: 10.1109/BigData.2016.7840915.
[7] R. A. Marcum, C. H. Davis, G. J. Scott, and T. W. Nivin, “Rapid broad area search and detection of Chinese surface-to-air missile sites using deep convolutional neural networks,” J. Appl. Remote Sens., vol. 11, no. 4, 2017, Art. no. 046214, doi: 10.1117/1.JRS.11.046214.
[8] A. B. Cannaday II, C. H. Davis, G. J. Scott, B. Rupprecht, and D. T. Anderson, “Broad area search and detection of surface-to-air missile sites using spatial fusion of component object detections from deep neural networks,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 13, pp. 4728–4737, 2020, doi: 10.1109/JSTARS.2020.3015662.
[9] J. Redmon and A. Farhadi, “YOLOv3: An incremental improvement,” 2018, arXiv:1804.02767v1. [Online]. Available: https://arxiv.org/abs/1804.02767v1.
[10] P. Gong et al., “Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017,” Sci. Bull., vol. 64, no. 6, pp. 370–373, Mar. 2019, doi: 10.1007/s11207-019-01862-9.
[11] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.
[54] G. Lin, A. Milan, C. Shen, and I. Reid, “RefineNet: Multi-path refinement networks for high-resolution semantic segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 1925–1934, doi: 10.1109/CVPR.2017.549.

[55] H. Wu, J. Zhang, K. Huang, K. Liang, and Y. Yu, “FastFCN: Rethinking dilated convolution in the backbone for semantic segmentation,” 2019, arXiv:1903.11816. [Online]. Available: http://arxiv.org/abs/1903.11816

[56] J. Chen, A. Liao, J. Chen, S. Peng, L. Chen, and H. Zhang, “30-meter global land cover data product—GlobeLand30,” Geomatics World, vol. 24, no. 1, pp. 1–8, 2017.

[57] X. Zhang, L. Liu, X. Chen, Y. Gao, and J. Mi, “GLC_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery,” Earth Syst. Sci. Data, vol. 13, no. 6, pp. 2753–2776, doi: 10.5194/essd-13-2753-2021.

[58] M. Haklay and P. Weber, “OpenStreetMap: User-generated street maps,” IEEE Pervasive Comput., vol. 7, no. 4, pp. 12–18, Oct.–Dec. 2008, doi: 10.1109/MPRV.2008.80.

[59] J. Chen, A. Liao, J. Chen, S. Peng, L. Chen, and H. Zhang, “30-meter global land cover data product—GlobeLand30,” Geomatics World, vol. 24, no. 1, pp. 1–8, 2017.

[60] M. Jing et al., “Detecting unknown dams from high-resolution remote sensing images: A deep learning and spatial analysis approach,” Int. J. Appl. Earth Observ. Geoinf., vol. 104, Dec. 2021, Art. no. 102576, doi: 10.1016/j.jag.2021.102576.

[61] X. Zhou, D. Wang, and P. Krähenbühl, “Objects as points,” 2019, arXiv:1904.07850. [Online]. Available: http://arxiv.org/abs/1904.07850

[62] Y. Hong, H. Pan, W. Sun, and Y. Jia, “Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes,” 2021, arXiv:2101.06085. [Online]. Available: http://arxiv.org/abs/2101.06085

[63] J. Glenn et al., “ultralytics/yolov5: V5.0 - YOLOv5-P6 1280 models, AWS, Supervise.ly and YouTube integrations (v5.0),” Zenodo, Apr. 2021, doi: 10.5281/zenodo.3908559.

Ning Li received the B.Sc. degree in geographic information science from Chang’an University, Xi’an, China, in 2018. He is currently working toward the Ph.D. degree in geography from Nanjing University, Nanjing, China.

His research interests include deep learning and remote sensing applications.

Liang Cheng received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2008.

He is currently a Professor with the Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing, China. His current research focuses on the integration of multsource spatial data and remote sensing image processing.

Chen Ji received the B.Eng. degree in surveying and mapping from Nanjing Normal University, Nanjing, China, in 2013, and the M.E. degree in geographic information systems from Beijing Normal University, Beijing, China, in 2016. He is currently working toward the Ph.D. degree in geography from Nanjing University, Nanjing, China.

His research interests include deep learning, GIS, and remote sensing applications.

Shengkun Dongye received the B.Eng. degree in science and technology of remote sensing from the Anhui University of Science and Technology, Huainan, China, in 2018, and the M.E. degree in surveying and mapping engineering from the Kunming University of Science and Technology, Kunming, China, in 2022. He is currently working toward the Ph.D. degree in geography from Nanjing University, Nanjing, China.

His research interests include remote sensing applications, deep learning, and GIS.

Manchun Li received the Ph.D. degree in cartography from Nanjing University, Nanjing, China, in 1992.

He is currently a Professor with the Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University. His research interests include GIS and remote sensing applications.