



RS IMAGE INTERPRETATION FROM A DATA PERSPECTIVE

Diversity, Richness, Scalability (DiRS) :

On Benchmarking Remote Sensing Image Interpretation

Gui-Song Xia

guisong.xia@whu.edu.cn

School of Computer Science, Wuhan University

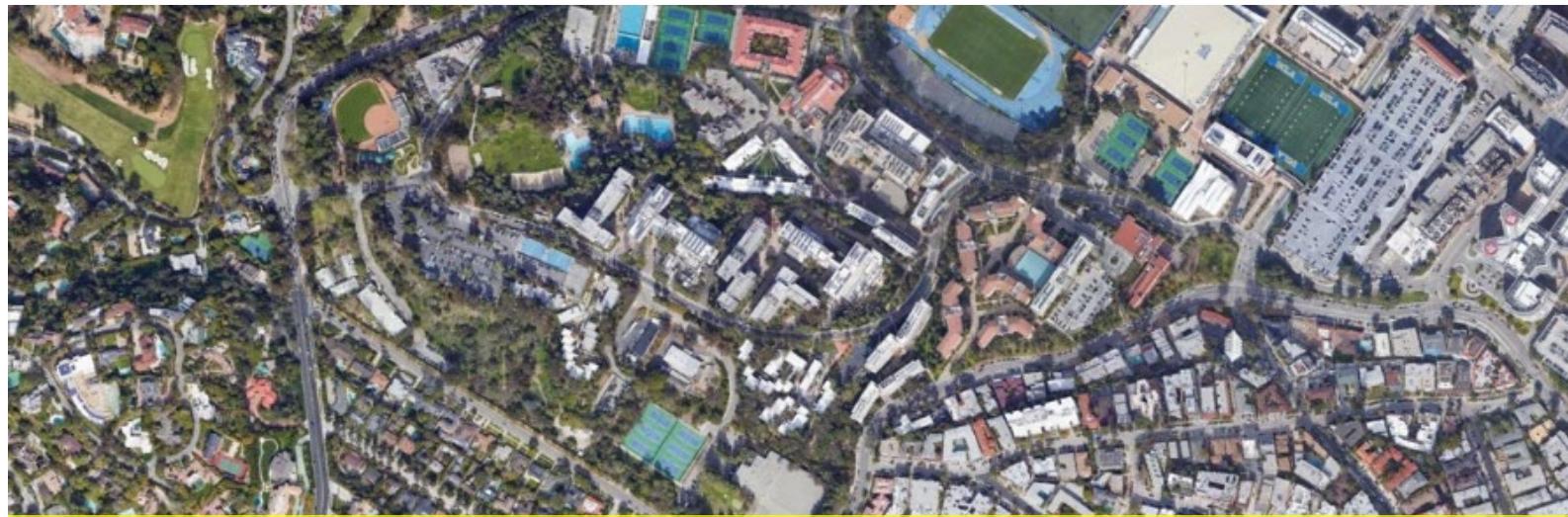
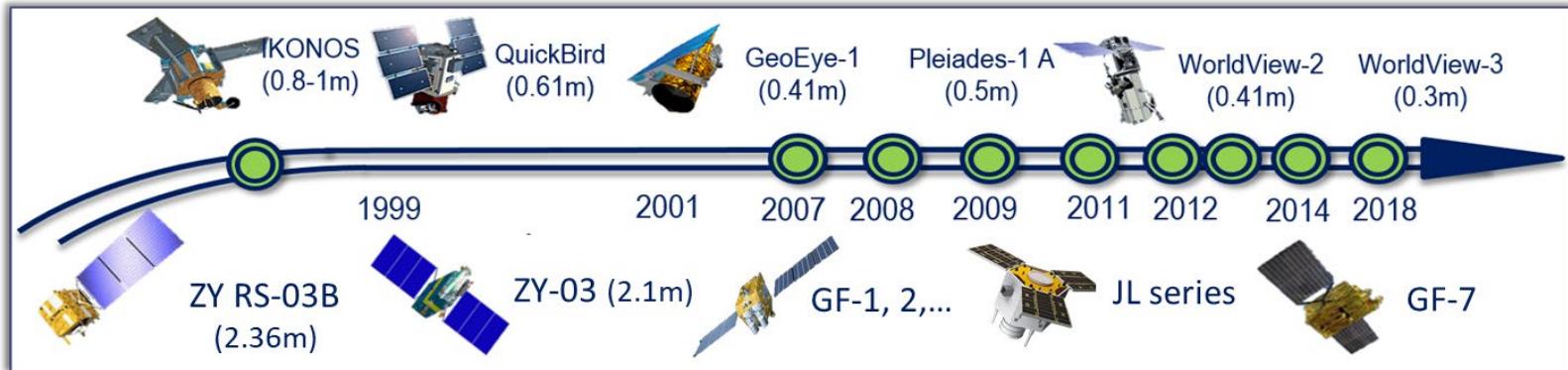
Institute of Artificial Intelligence, Wuhan University

State Key Lab. LIESMARS, Wuhan University

Jun. 1, 2020

Advanced RS Technology

RS technology has significantly improved the earth observation ability.



The characterization of features on the earth surface.

Applications of RS Images

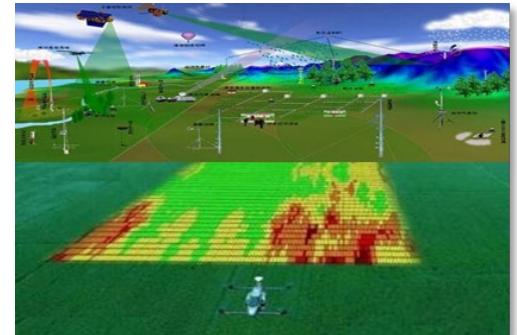
Interpretation of RS images plays important roles in many real-world applications.



National security



High definition map



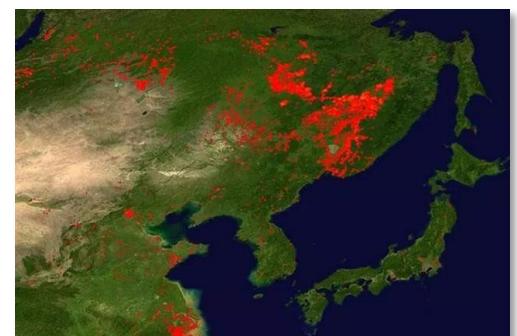
Precision agriculture



Smart city



Disaster assessment



Environ. monitoring

Interpretation of RS Images

Current situation: Increasing demands for automatic interpretation



Satellites on-orbit

- **Variation:** difference in spectral, spatial, and temporal properties

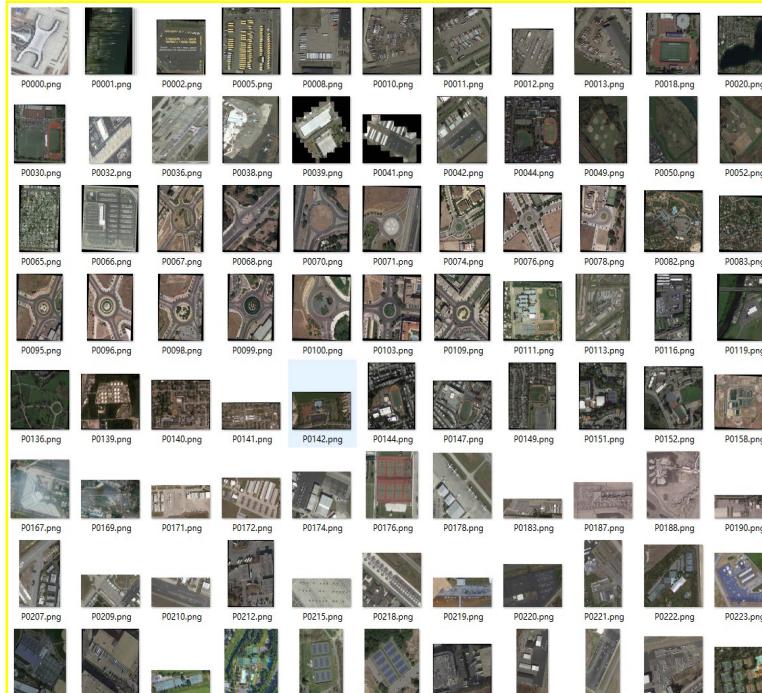


Image acquisition

- **Inconsistency:** multi-modal, multi-source RS images

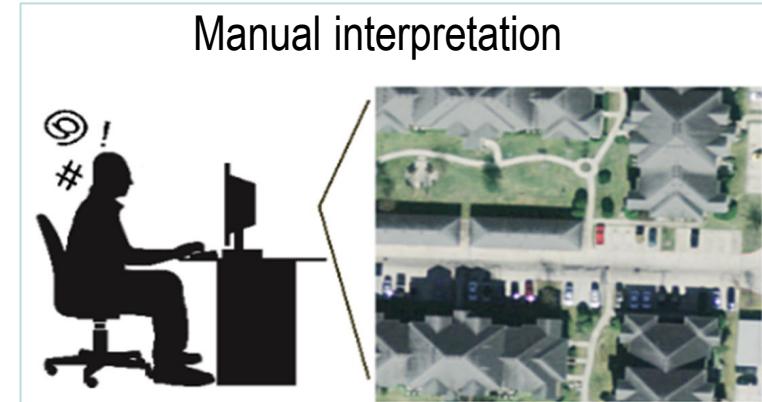
Interpretation of RS Images

Current situation: Increasing demands for automatic interpretation



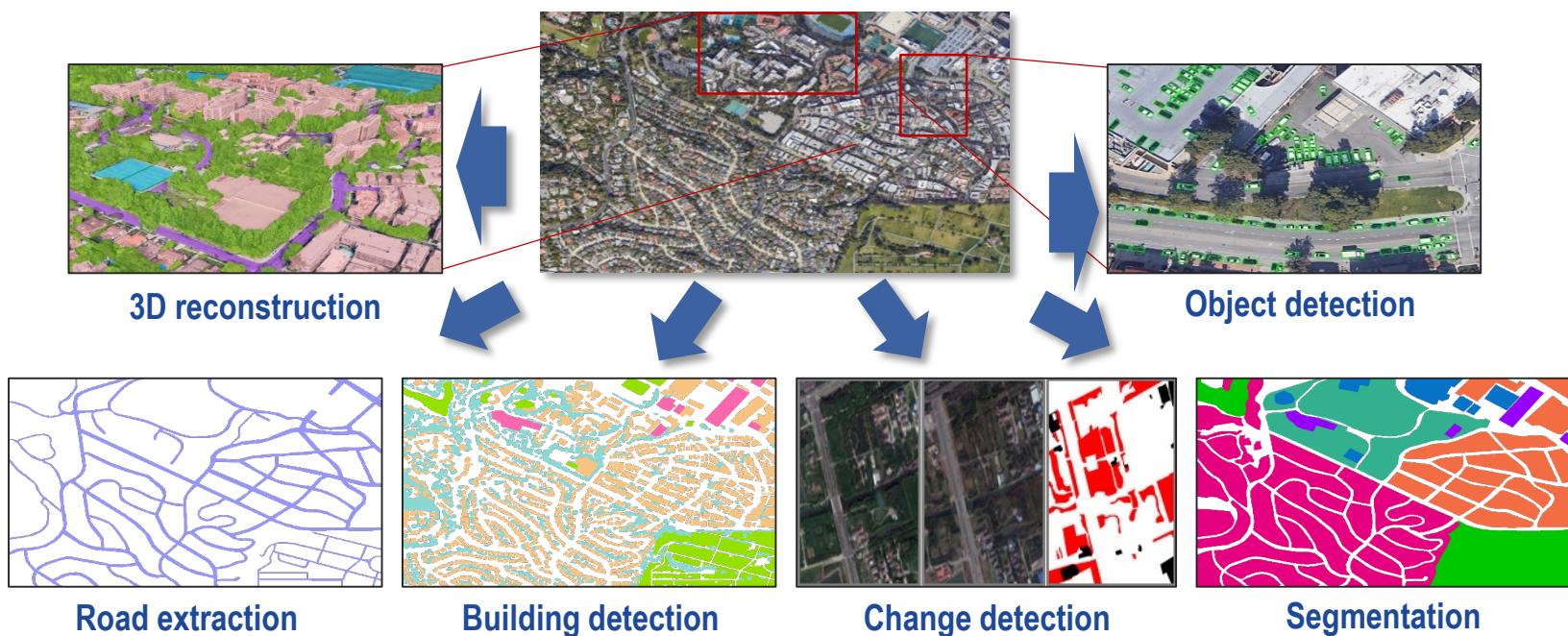
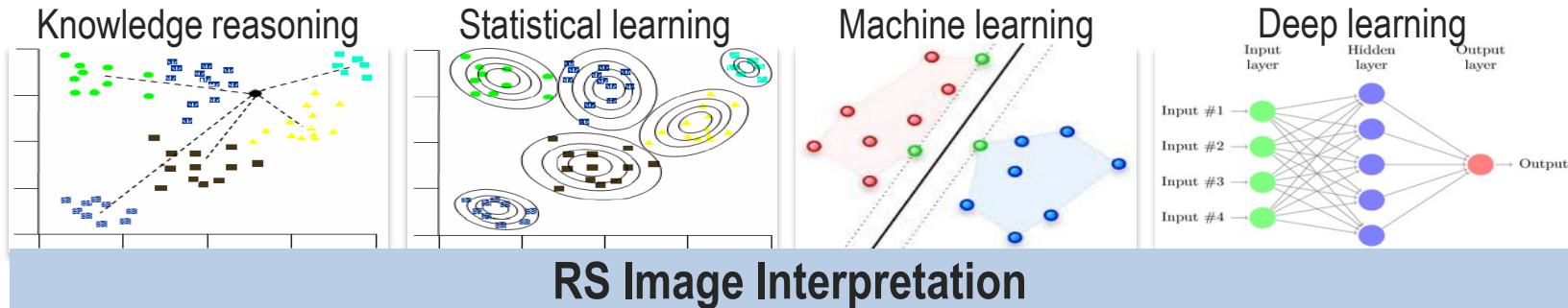
Large volume of images

- **Challenge:** geometrical shapes, textural attribute, structural characteristic ...



Blooming Data-driven methods

Content interpretation: data-driven methods for RS image interpretation.



- Huge-volume RS images **v.s.** *limited data with labels*
- Increasing number of datasets with *different purposes and standards*
 - The **ever-growing volume of RS images** is acquired while very **few of them are annotated** with valuable information.
 - The **generalization ability of algorithms** for interpreting RS images is of great urgency to be enhanced.
 - The **Representative and large-scale RS image datasets** with accurate annotations is demanded to narrow *the gap between algorithm development and real applications*.
 - There is a **lack of public platforms** for systematic evaluation and fair comparison among different algorithms.



Outline

- Background
- **Research Focus in the Past Decade**
- Guidances to Benchmark RS Image Interpretation
- An Example: Million-AID
- Challenges and Perspectives
- Conclusions

Focus in the Past Decade



■ A systematic investigation to the literature

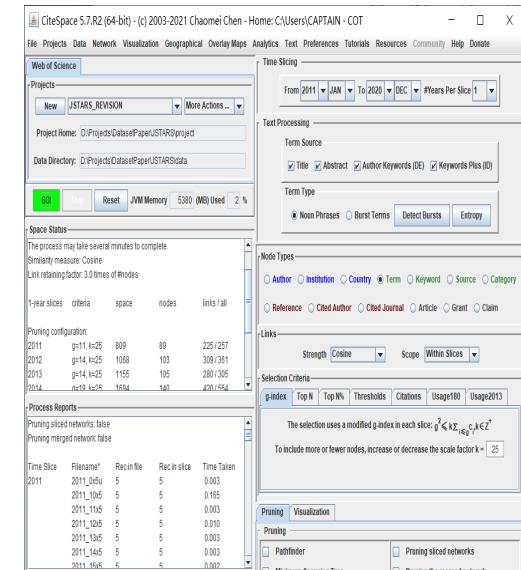
- Journals with good reputation: *ISPRS J. P&RS, RSE, TGRS ...*
- Meta-data for analysis: 5, 827 surveyed articles over the past decade
- Bibliometric analysis: title/topic/keywords ... concerning image interpretation



Selected journals



Meta-data

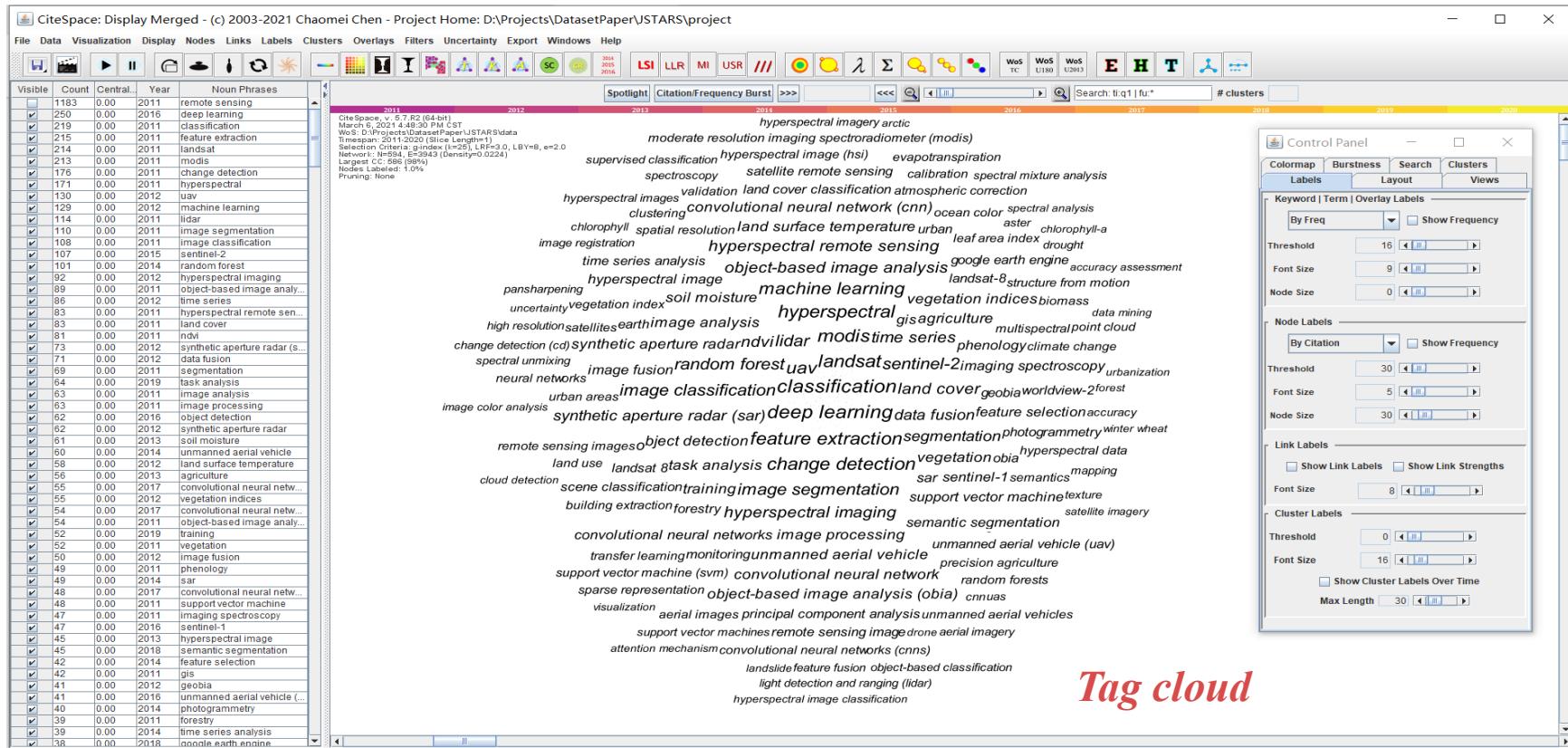


Bibliometric analysis

Frequency Terms



- Interpretation mainly focus on **classification** tasks (scene, land cover, ...)
 - **Change detection**, **segmentation**, and object detection occupy prominent positions
 - **Deep learning** and **feature extraction** play significant roles in RS image interpretation



Tag cloud

Available Datasets for Interpretation

■ RS *scene classification* datasets

Dataset	#Cat.	#Images per cat.	#Instances	Resolution (m)	Image size	GL/IT/SP	Year
UC-Merced	21	100	2,100	0.3	256×256	XXXX	2010
WHU-RS19	19	50 to 61	1,013	up to 0.5	600×600	XXXX	2012
RSSCN7	7	400	2,800	—	400×400	XXXX	2015
SAT-4	4	89,963 to 178,034	500,000	1 to 6	28×28	XXXX	2015
SAT-6	6	10,262 to 150,400	405,000	1 to 6	28×28	XXX	2015
BCS	2	1,438	2,876	—	600×600	XX✓	2015
RSC11	11	~100	1,232	~0.2	512×512	XXXX	2016
SIRI-WHU	12	200	2,400	2	200×200	XXXX	2016
NWPU-RESISC45	45	700	31,500	0.2 to 30	256×256	XXXX	2016
AID	30	220 to 420	10,000	0.5 to 8	600×600	XXXX	2017
RSI-CB256	35	198 to 1,331	24,000	0.3 to 3	256×256	XXXX	2017
RSI-CB128	45	173 to 1,550	36,000	0.3 to 3	128×128	XXXX	2017
Planet-UAS	17	—	40,480	3 to 5	256×256	✓✓✓	2017
RSD46-WHU	46	500 to 3,000	117,000	0.5 to 2	256×256	XXX	2017
MASATI	7	304 to 1,789	7,389	—	512×512	XXXX	2018
EuroSAT	10	2,000 to 3,000	27,000	10	64×64	✓✓✓	2018
PatternNet	38	800	30,400	0.06 to 4.7	256×256	XXX	2018
fMoW	62	—	132,716	0.5	74×58 to 16,184×16,288	✓✓✓	2018
WiDS Datathon 2019	2	—	20,000	3	256×256	XXX	2019
Optimal-31	31	60	1,860	—	256×256	XXX	2019
BigEarthNet	43	328 to 217,119	590,326	10,20,60	20×20;60×60;120×120	✓✓✓	2019
CLRS	25	600	15,000	0.26 to 8.85	256×256	XXX	2020
MLRSN	46	1,500 to 3,000	109,161	0.1 to 10	256×256	XXX	2020

■ RS *object detection* datasets

Datasets	Annot.	#Cat.	#Instances	#Images	Resolution (m)	Image width	GL/IT/SP	Year
TAS	HBB	1	1,319	30	—	792	XXXX	2008
OIRDS	OBJ	5	1,800	900	up to 0.08	256 to 640	✓✓✓	2009
SZTAKI-INRIA	OBJ	1	665	9	—	~800	XXXX	2012
NWPU-VHR10	HBB	10	3,651	800	0.08 to 2	~1,000	XXXX	2014
DLR-MVDA	OBJ	2	14,235	20	0.13	5,616	XX✓	2015
UCAS-AOD	OBJ	2	14,596	1,510	—	~1,000	XXXX	2015
VEDAI	OBJ	9	3,640	1,210	0.125	512;1,024	✓XX	2016
COWC	CP	1	32,716	53	0.15	2,000 to 19,000	✓XX	2016
HRSC2016	OBJ	26	2,976	1,061	—	~1,100	XXXX	2016
RSOD	HBB	4	6,950	976	0.3 to 3	~1,000	XXXX	2017
CARPK	HBB	1	89,777	1,448	—	1,280	XX✓	2017
SSDD/SSDD+	HBB/OBB	1	2,456	1,160	1 to 15	~500	XX✓	2017
SpaceNet1-6*	Polygon	1	859,982	—	up to 0.3	—	✓✓✓	2018
LEVIR	HBB	3	11,028	22,000	0.2 to 1	800	XXX	2018
VisDrone	HBB	10	54,200	10,209	—	2,000	XXX	2018
xView	HBB	60	1,000,000	1,413	0.3	~3,000	✓✓✓	2018
DOTA-v1.0	OBJ	15	188,282	2,806	up to 0.3	800 to 13,000	XXXX	2018
ITCV	HBB	1	29,088	173	0.1	3,744;5,616	XXXX	2018
WHU building dataset	Polygon	1	221,107	25,420	0.075 to 2.7	512	XXXX	2018
DeepGlobe Building	Polygon	2	302,701	24,586	0.3	650	XX✓	2018
OpenSARShip	Clip	1	11,346	41	~10	—	✓✓✓	2018
CrowdAI Mapping Challenge	Polygon	1	2,910,917	341,058	—	300	XXX	2018
Airbus Ship Detection Challenge	Polygon	1	~131,000	208,162	—	768	XXX	2018
iSAID	Polygon	15	655,451	2,806	up to 0.3	800 to 4,000	XXXX	2019
HRRSD	HBB	13	55,740	21,761	0.15 to 1.2	152 to 10,569	XXXX	2019
DIOR	HBB	20	192,472	23,463	0.5 to 30	800	XXXX	2019
DOTA-v1.5	OBJ	16	402,089	2,806	up to 0.3	800 to 13,000	XXXX	2019
SAR-Ship-Dataset	HBB	1	5,9535	43,819	up to 3	256	XX✓	2019
AIR-SARShip	HBB	1	2,040	300	1:3	1,000	✓✓✓	2020
HRSID	HBB	1	16,951	5,604	0.5;1:3	800	XX✓	2020
RarePlanes	Polygon	1	644,258	50,253	0.3	—	✓✓✓	2020
DOTA-v2.0	OBJ	18	1,793,658	11,268	up to 0.3	800 to 20,000	XXX	2020

Available Datasets for Interpretation



RS semantic segmentation datasets

Datasets	#Cat.	#Images	Resolution (m)	#Channels	Image size	GL/IT/SP	Year
Kennedy Space Center	13	1	18	224	512×614	✓✓✓	2005
Botswana	14	1	30	242	1,476×256	✓✓✓	2005
Salinas	16	1	3.7	224	512×217	✓✓✓	—
University of Pavia	9	1	1.3	115	610×340	✓✓✓	—
Pavia Centre	9	1	1.3	115 bands	1,096×492	✓✓✓	—
ISPRS Vaihingen	6	33	0.09	IR, R, G, DSM, nDSM	~2,500×2,500	✓✓✓	2012
ISPRS Potsdam	6	38	0.05	IR, RGB, DSM, nDSM	6,000×6,000	✓✓✓	2012
Massachusetts Buildings	2	151	1	RGB	1,500×1,500	✓✓✓	2013
Massachusetts Roads	2	1,171	1	RGB	1,500×1,500	✓✓✓	2013
Indian Pines	16	1	20	224	145×145	✓✓✓	2015
Zurich Summer	8	20	0.62	NIR, RGB	1,000×1,000	✓✓✓	2015
SPARCS Validation	8	80	30	11	1,000×1,000	✓✓✓	2016
Biome	4	96	30	11	~9,000×9,000	✓✓✓	2017
Inria	2	360	0.3	RGB	5,000×5,000	✓✓✓	2017
EvLab-SS	10	60	0.1 to 2	RGB	4,500×4,500	✓✓✓	2017
RIT-18	18	3	0.047	6	9,000×6,000	✓✓✓	2017
CITY-OSM	3	1,671	0.1	RGB	2,500×2,500 to 3,300×3,300	✓✓✓	2017
Dstl-SIFD*	10	57	up to 0.3	up to 16	~3,350×3,400	✓✓✓	2017
IEEE GRSS Data Fusion Contest 2017	17	30	1.4	9	643×666; 374×515	✓✓✓	2017
IEEE GRSS Data Fusion Contest 2018	20	1	1	48	4,172×1,202	✓✓✓	2018
Aeroscapes	11	3,269	—	RGB	720×1,280	✓✓✓	2018
DLRSD	17	2,100	0.3	RGB	256×256	✓✓✓	2018
DeepGlobe Land Cover	7	1,116	0.5	RGB	2,448×2,448	✓✓✓	2018
S2Globe-LCZ42	17	400,673	10	10	32×32	✓✓✓	2019
SEN12MS	33	180,662 triplets	10 to 50	up to 13	256×256	✓✓✓	2019
95-Cloud	1	43,902	30	NIR, RGB	384×384	✓✓✓	2019
Shakeel et al.	1	2,682	0.3	RGB	300×300	✓✓✓	2019
ALCD Cloud Masks	8	38	10	RGB	1,830×1,830	✓✓✓	2019
SkyScapes	31	16	0.13	RGB	5,616×3,744	✓✓✓	2019
DroneDeploy	7	55	0.1	RGB	up to 12,039×13,854	✓✓✓	2019
Slovenia LULC	10	940	10	6	5,000×5,000	✓✓✓	2019
LandCoverNet	7	1,980	10	NIR, RGB	256×256	✓✓✓	2020
UAvid	8	420	—	RGB	~4,000×2,160	✓✓✓	2020
GID	15	150	0.8 to 10	4	6,800×7,000	✓✓✓	2020
LandCover.ai	3	41	0.25 to 0.5	RGB	9,000×9,500; 4,200×4,700	✓✓✓	2020
Agriculture-Vision	9	94,986	0.1; 0.15; 0.2	NIR, RGB	512×512	✓✓✓	2020
S2CMC*	18	513	20	13	1,024×1,024	✓✓✓	2020

RS change detection datasets

Datasets	#Cat.	#Image pairs	Resolution (m)	#Channels	Image size	GL/IT/SP	Year
SZTAKI AirChange	2	13	1.5	RGB	952×640	✓✓×	2009
AICD	2	1,000	0.5	115	800×600	✗✗✗	2011
Taizhou Data	4	1	30	6	400×400	✓✓✓	2014
Kunshan Data	3	1	30	6	800×800	✓✓✓	2014
Cross-sensor Bastrop	2	4	30,120	7,9	444×300; 1,534×808	✓✓✓	2015
MtS-WH	9	1	1	NIR, RGB	7,200×6,000	✓✓✓	2017
Yancheng	4	2	30	242	400×145	✓✓✓	2018
GETNET dataset	2	1	30	198	463×241	✗✓✓	2018
Urban-rural boundary of Wuhan	20	1	4/30	4, 9	960×960	✓✓✓	2018
Hermiston City, Oregon	5	1	30	242	390×200	✓✓✓	2018
OSCD	2	24	10	13	600×600	✓✓✓	2018
WHU building dataset	2	1	0.2	RGB	32,507×15,354	✓✓✓	2018
Season-varing dataset	2	16,000	0.03 to 0.1	RGB	256×256	✗✗✗	2018
ABCD	2	16,950	0.4	RGB	128×128; 160×160	✗✓✗	2018
California flood dataset	2	1	5.30	RGB, 11	1534×808	✓✓✓	2019
López-Fandiño et al.	5	2	20	224	984×740; 600×500	✓✓✓	2019
xBD	6	11,034	up to 0.8	RGB	1,024×1,024	✓✓✓	2019
HRSCD	6	291	0.5	RGB	10,000×10,000	✓✓✓	2019
LEVIR-CD	2	637	0.5	RGB	1,024×1,024	✗✗✗	2020
SECOND	30	4,214	0.5 to 3	RGB	512×512	✗✗✗	2020
Google Dataset	2	1,067	0.55	RGB	256×256	✓✓✓	2020
Zhang et al.	2	4	2;2.4;5.8	NIR, RGB	1,431×1,431; 458×559; 1,154×740	✓✓✓	2020
Hi-UCD	9	1,293	0.1	RGB	1,024×1,024	—/—/Y	2020
SpaceNet7	—	24	4	RGB	—	✓✓✓	2020
S2MTCP	2	1,520	up to 10	13	600×600	✓✓✓	2021

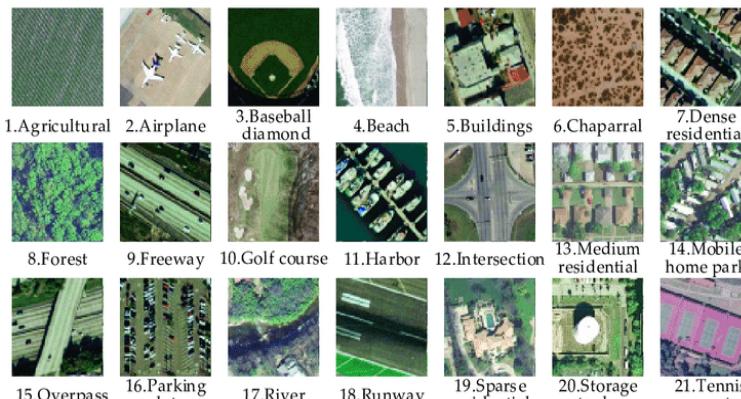
Some Critical Reviews

■ Categories involved in interpretation

- **Small number** of categories, content interpretation for **certain objects**
- Categories with **equal relationship**, chaotic management for semantic information
- Complex semantic categories and relationships in real applications, e.g., LULC

■ Dataset annotation

- Nearly all **manually annotated** by experts, extensive **labor remains to relieve**
- **Visualization for large scale, high spectral RS images** annotation is demanded
- **Lack of interchange** with application departments for efficient data annotation



UC Merced, 21 classes, 2100 images



NPWU VHR, 800 images, manually annotated

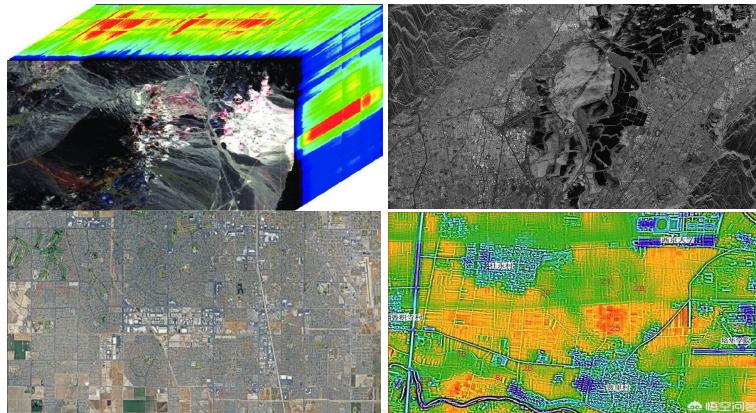
Some Critical Reviews

■ Image source

- Optical images (Google Earth) as data standard since spatial pattern, visual texture, structural information are more concerned (e.g., for scene/object recognition)
- High spectral, SAR images for abnormal object detection by the physical property

■ Dataset scale

- *Limited number, size of chipped images, performance saturation* of algorithms
- *Lack of image variation, sample diversity*, and *content representation*, causing *weak generalization ability* of interpretation algorithms



Multi-modal image source



Simple scenes and complex reality



Outline

- Background
- Research Focus in the Past Decade
- **Guidances to Benchmark RS Image Interpretation**
- An Example: Million-AID
- Challenges and Perspectives
- Conclusions

Guidances to Benchmark RS II

■ Toward real-world scenarios *rather than specific algorithms*

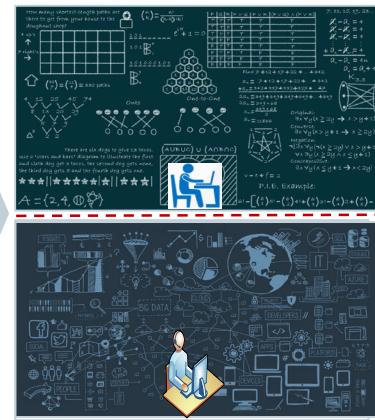
- Model training, testing, and screening for practical applications
- Rich samples of variation in background, scale, imaging conditions, ...

■ Annotation by application sides *rather than algorithm developers*

- Label images and samples considering practical challenges in application

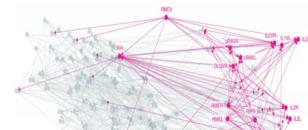


Images



Annotators

Algorithm designer



Algorithm oriented

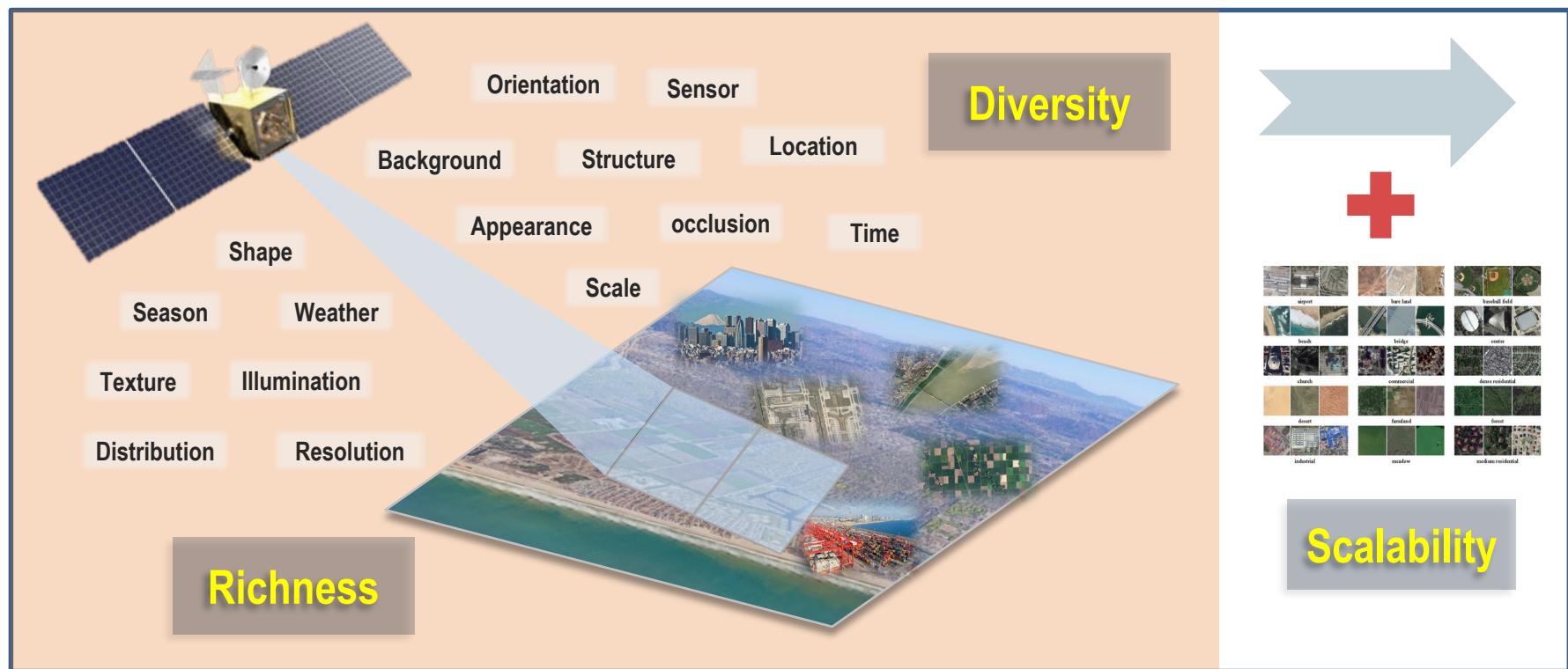


Application oriented



■ DiRS for dataset construction

- **Diversity**: between-/within-class diversity, complementarity of features
- **Richness**: large-scale images, sufficient samples, diverse characteristics ...
- **Scalability**: sufficient space for new data involvement, sustainable availability

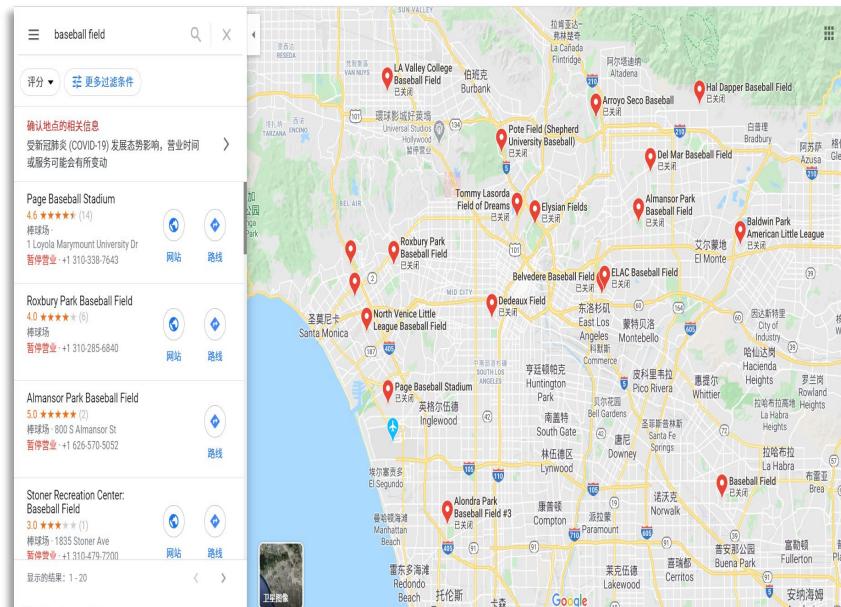


Geographic Information Integration

■ Coordinates Collection for RS Image Acquisition

■ Geographic information utilization

- Rich positional data with millions of point, line and region objects
- Inherent semantic tags for interested image extraction with Map API



The screenshot shows a Google Maps search interface for 'baseball field'. The search bar at the top contains the query 'baseball field'. Below the search bar, there is a sidebar with filtering options: '更多过滤条件' (More filtering conditions), '受新冠影响 (COVID-19) 发展态势影响, 营业时间或服务可能会有所变动' (Affected by COVID-19 development trend, business hours or services may change), and three specific results listed:

- Page Baseball Stadium**: 4.6 stars (14 reviews). Address: 1 Loyola Marymount University Dr,暫停營業 +1 310-335-7643. Buttons: 網站 (Website) and 路線 (Directions).
- Roxbury Park Baseball Field**: 4.0 stars (6 reviews). Address: 諸停營業 +1 310-285-6840. Buttons: 網站 (Website) and 路線 (Directions).
- Almansor Park Baseball Field**: 5.0 stars (2 reviews). Address: 800 S Almansor St,暫停營業 +1 626-570-5052. Buttons: 網站 (Website) and 路線 (Directions).

The main map view shows numerous baseball fields marked with red location pins across the Los Angeles area, including Sun Valley, Burbank, Hollywood, and surrounding neighborhoods.



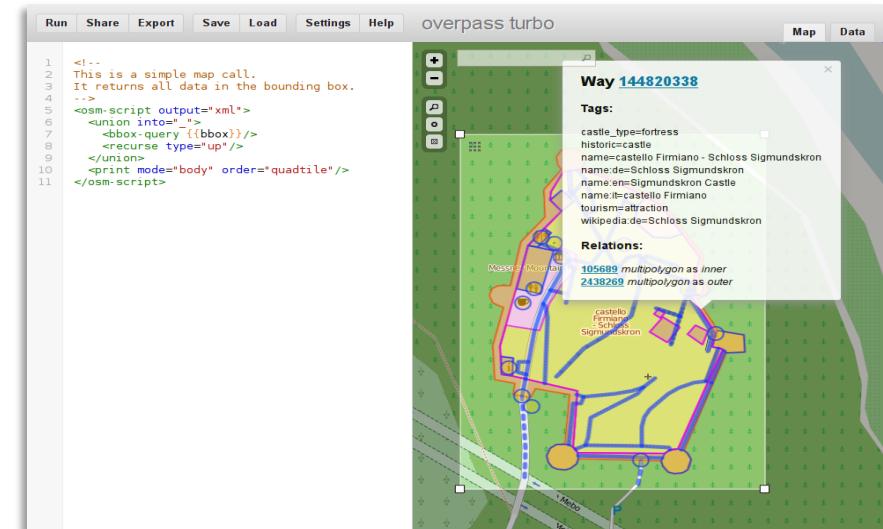
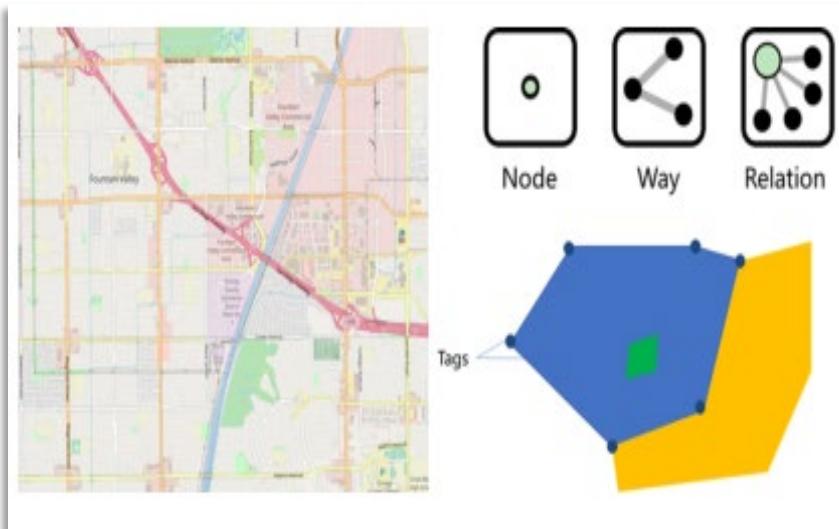
Searched baseball fields using Google Map API

Geographic Information Integration

■ Coordinates Collection for RS Image Acquisition

■ Open source data

- Geographic data with rich semantic information that is timely updated, low cost and with large amount, e.g., OSM, WikiMapia ...
- Excellent interface for data customization, elements aligned with different maps



The screenshot shows the overpass turbo web application interface. At the top, there are buttons for Run, Share, Export, Save, Load, Settings, Help, and a search bar. The main area has tabs for Map and Data. On the left, a code editor displays an Overpass query:

```
1  <!--
2  This is a simple map call.
3  It returns all data in the bounding box.
4  -->
5  <osm-script output="xml">
6  <union into="">
7  <bbox-query {{bbox}}/>
8  <recurse type="up"/>
9  </union>
10 <out mode="body" order="quadtile"/>
11 </osm-script>
```

On the right, a map view shows a castle area with various features highlighted in pink and purple. A tooltip for a way is displayed, providing details such as:

Way 144820338

Tags:

- castle
- fortress
- historical
- name=castello Firmiano - Schloss Sigmundskron
- name=desSchloss Sigmundskron
- name=en:Sigmundskron Castle
- name=it:castello Firmiano
- tourism=attraction
- wikipedia.de=Schloss Sigmundskron

Relations:

- 105699 multipolygon as inner
- 2438269 multipolygon as outer

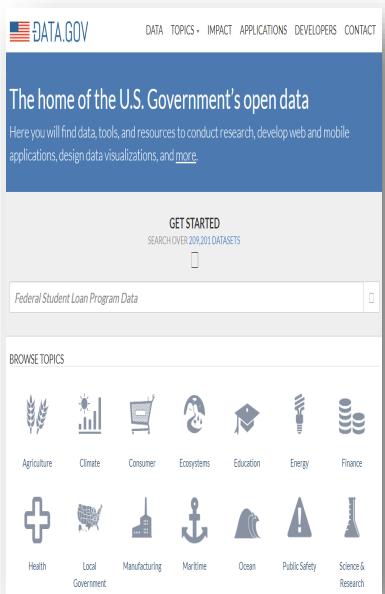
Elements of interest extracted from OSM

Geographic Information Integration

■ Coordinates Collection for RS Image Acquisition

■ Geodatabase integration

- Public geodatabases released by the state institutions and communities
- Domain-specific geodatabase that is publicly available



The home of the U.S. Government's open data. Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and more.

GET STARTED
SEARCH OVER 209,201 DATASETS

Federal Student Loan Program Data

BROWSE TOPICS

- Agriculture
- Climate
- Consumer
- Ecosystems
- Education
- Energy
- Finance
- Health
- Local Government
- Manufacturing
- Maritime
- Ocean
- Public Safety
- Science & Research



Geospatial at the Bureau of Transportation Statistics

Search for open data

Data by Mode

- Rail
- Aviation
- Roads
- Transit
- Marine

Data by Category

- Performance
- Safety
- Freight
- Energy and Environment

Explore HIFLD Data

Find Data

Explore All Data



HIFLD Open Data

This site provides National foundation-level geospatial data within the open public domain that can be useful to support community preparedness, resilience, research, and more. The data is available for download as CSV, KML, Shapefile, and accessible via web services to support application development and data visualization.

For updates about HIFLD data, guidance on metadata, and known shapefile conversion issues, please access the HIFLD notifications page located [here](#).

Find Data

Search HIFLD Open

Explore All Data

National Bridge Inventory

Transit Stations

Interstate Highways

Water System

Protestant Churches

ESRI Open Data Hub

.....

Public geodatabases available for image coordinates collection

■ Manual Annotation

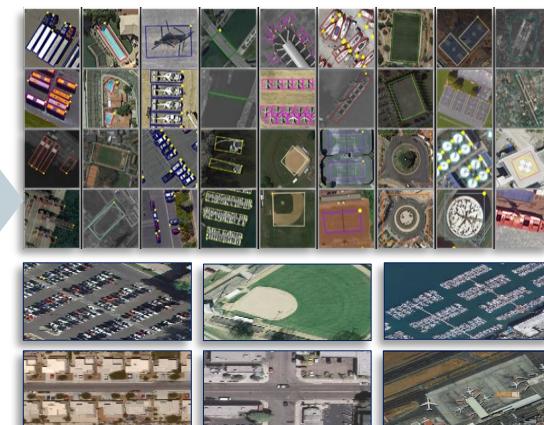
- Accuracy guarantee, but labor-intensive and time-consuming
- Hard to meet the scale requirements particularly for data-driven methods

■ Automatic Annotation

- Reduce the cost of annotation by leveraging learning models
- Bias problem deriving from the initialization data and model capability



Source image

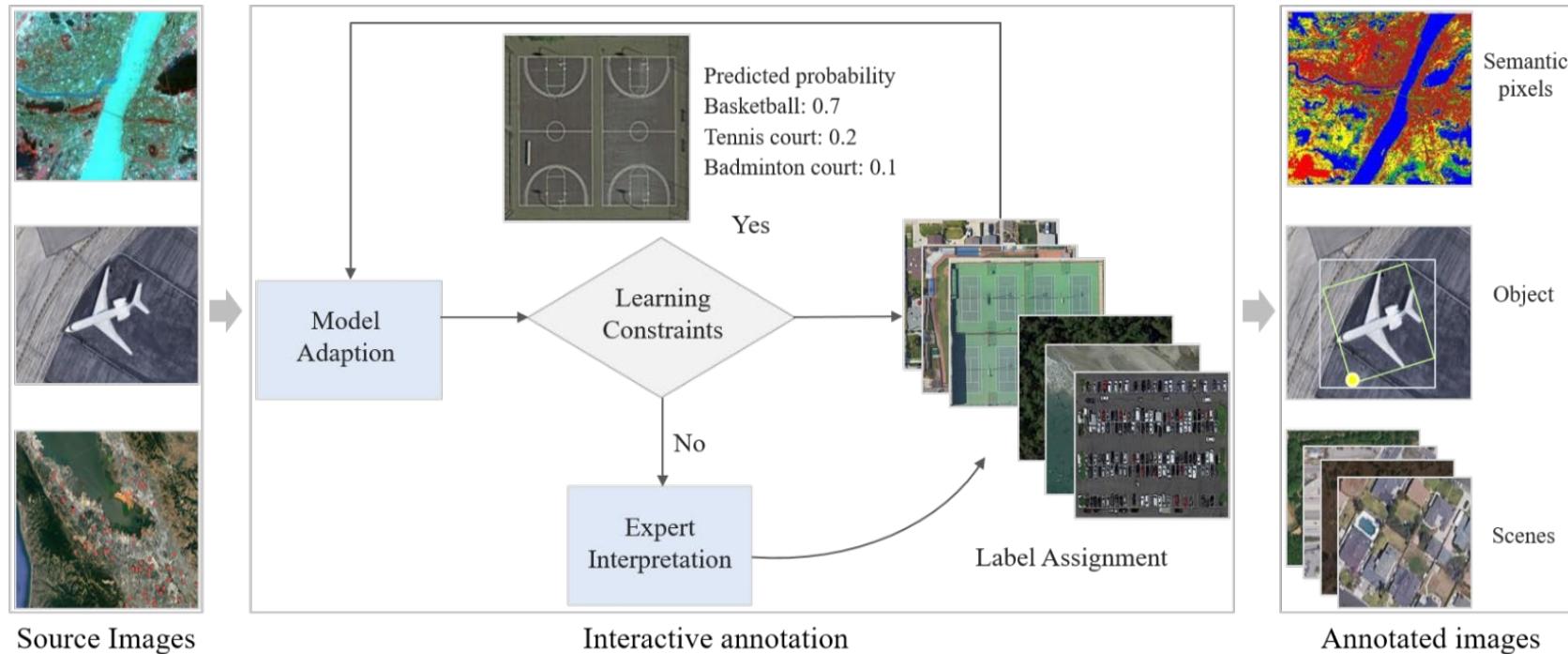


Annotated samples

Annotation methodology

■ Interactive Annotation

- Annotation with human-computer interaction, semi-automatic annotation
- Guarantee for quality and efficiency, toward large-scale dataset construction



General workflow of Semi-automatic annotation in RS images

Quality Assessment

- **Rules and Samples:** annotation without ambiguity, specific samples for instructions
- **Training of Annotators:** well-qualified annotators for dataset quality guarantee
- **Multi-stage Pipeline:** annotation task decomposition
- **Grading and Reward:** incentive mechanism for incompetent/competent annotators
- **Multiple Annotations:** merge multiple accurate annotations
- **Annotation Review:** expert/peer review and quality rating
- **Spot Check and Assessment:** gold data for annotation quality assurance

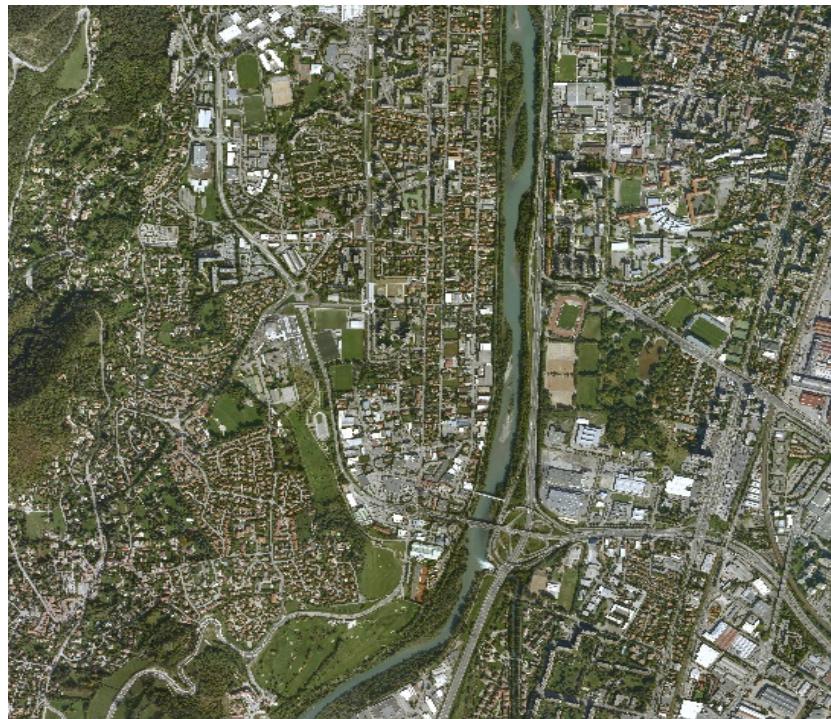


Outline

- Background
- Research Focus in the Past Decade
- Guidances to Benchmark RS Image Interpretation
- **An Example: Million-AID**
- Challenges and Perspectives
- Conclusions

■ Scene Classification

- High-level knowledge expression to RS image contents
- Semantic information recognition to local areas of RS images



Forest

Grass land

Parking lot

Resid. area

Indus. area

Water

Other

■ Category Organization

Chinese Land Use Classification Criteria



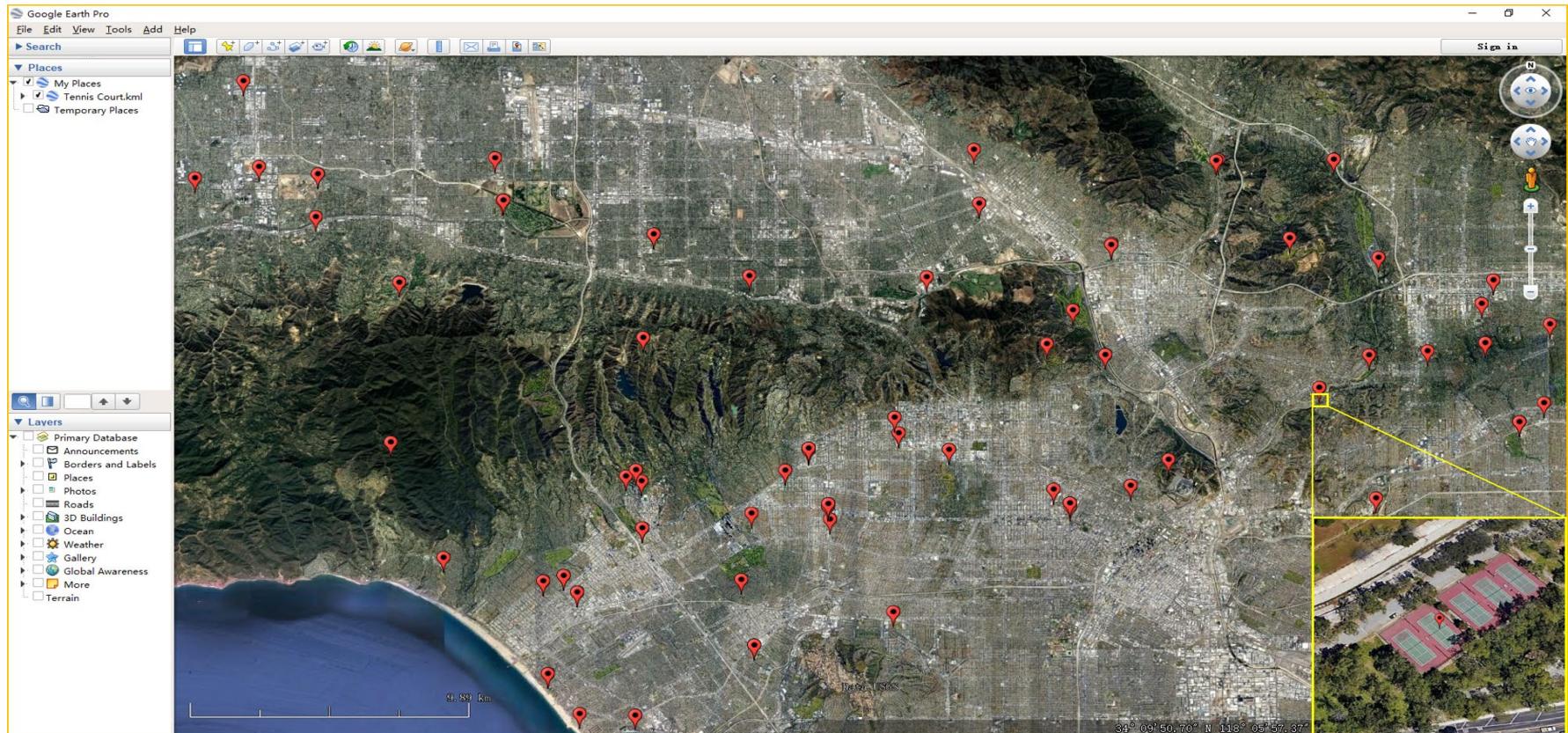
8 major categories with 51 sub-categories



The hierarchical scene category network of Million-AID

■ Semantic Coordinates Collection

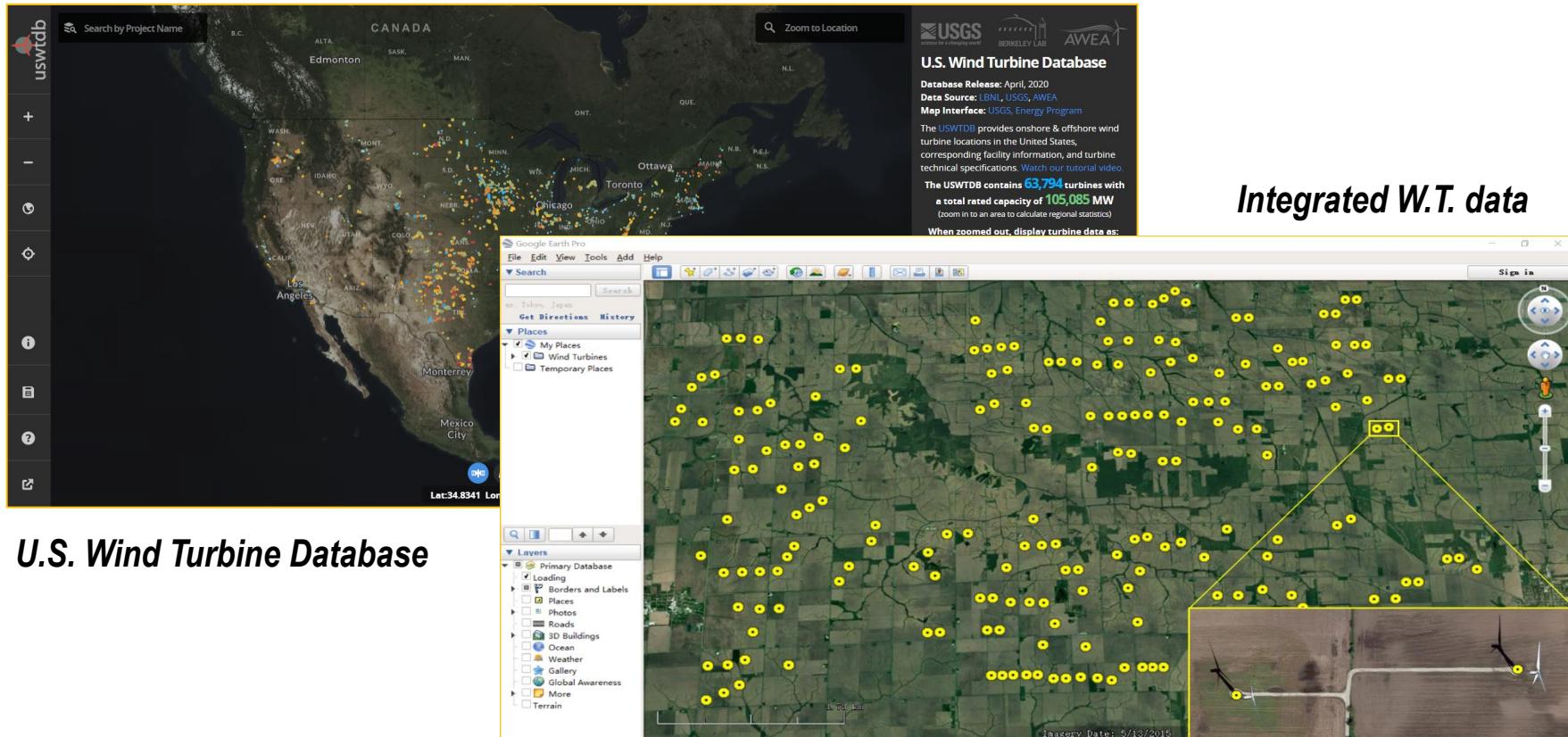
■ Point coordinates using Google Map API



The points of searched tennis courts shown in Google Earth. We consider the tennis courts as point ground features. In this case, we use Google Map API for coordinates collection.

Semantic Coordinates Collection

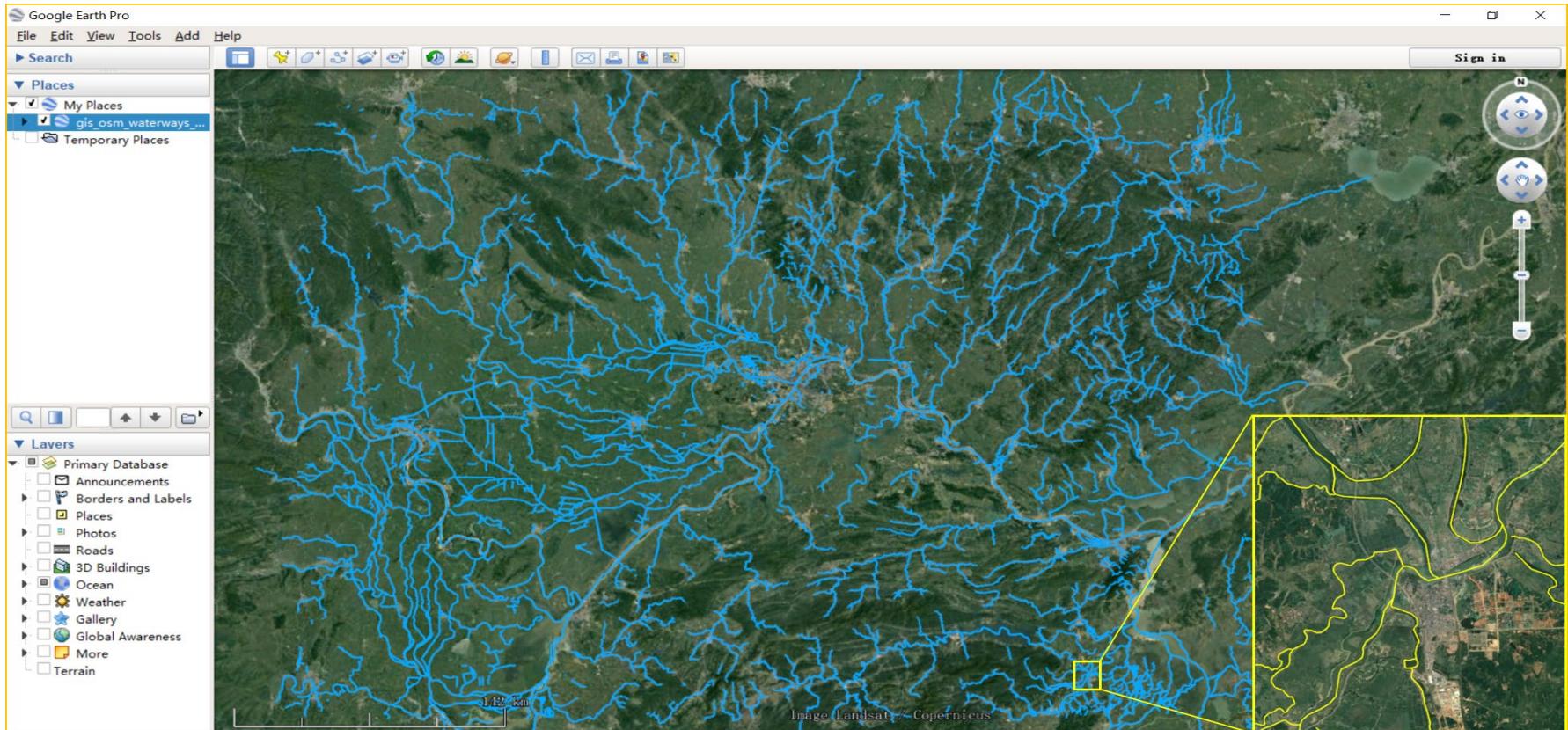
Point coordinates integrated from Geodatabase



The points of wind turbines extracted from USWTDB and integrated in Google Earth. Generally, over 60, 000 objects of wind turbines can be collected by the database.

■ Semantic Coordinates Collection

■ Line features extracted from OSM



The river lines within a local area of China collected from OSM and displayed in Google Earth

Semantic Coordinates Collection

Plane features customized on OSM

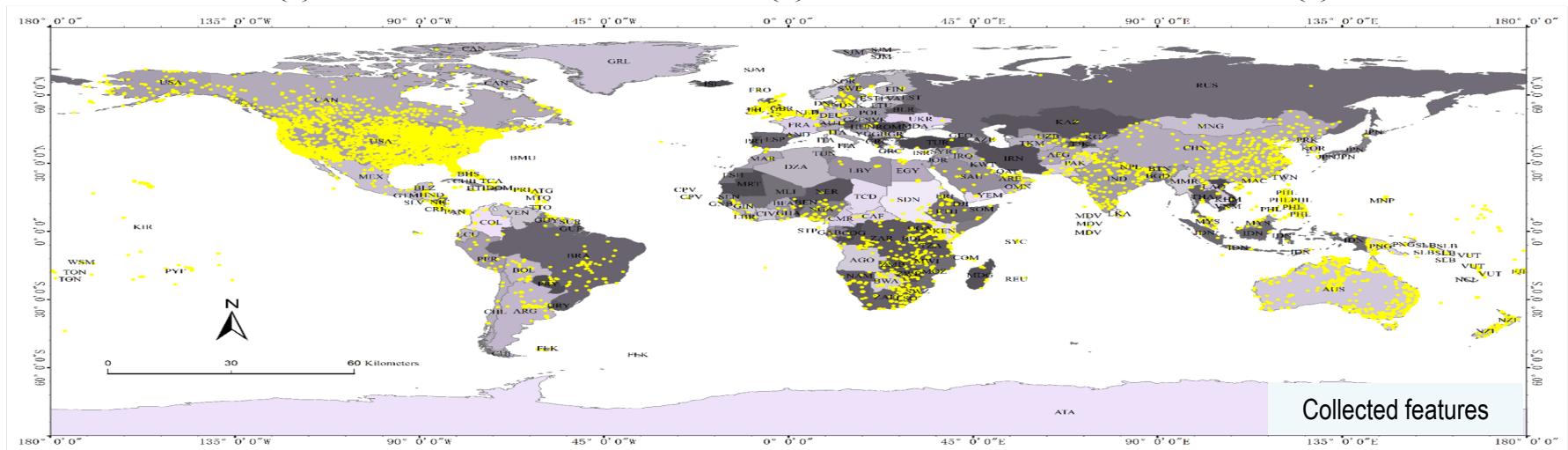


aeroway	aerodrome
ele	37
iata	VRC
icao	RPUV
length	1886
name	Virac Airport
operator	Civil Aviation Authority of the Philippines
source	Wikipedia
wikidata	
wikipedia	

(b) Semantic tags

```
<osm-script>
<union>
<query type="way">
<has-kv k="aeroway" regv="aerodrome"/>
<has-kv k="name" regv="[a|A]irport$|机场$"/>
</query>
<query type="relation">
<has-kv k="aeroway" regv="aerodrome"/>
<has-kv k="name" regv="[a|A]irport$|机场$"/>
</query>
<recurse type="relation-way"/>
</union>
<union>
<item/>
<recurse type="down"/>
</union>
<print/>
</osm-script>
```

(c) Data customization



The illustration of searching scenes of airports around the world. An airport in OSM contains a large amount of tags, which can be employed to search airports with specific semantic key-value labels.

■ Scene Image Acquisition

- **Image blocks by the line, point, and plane data**

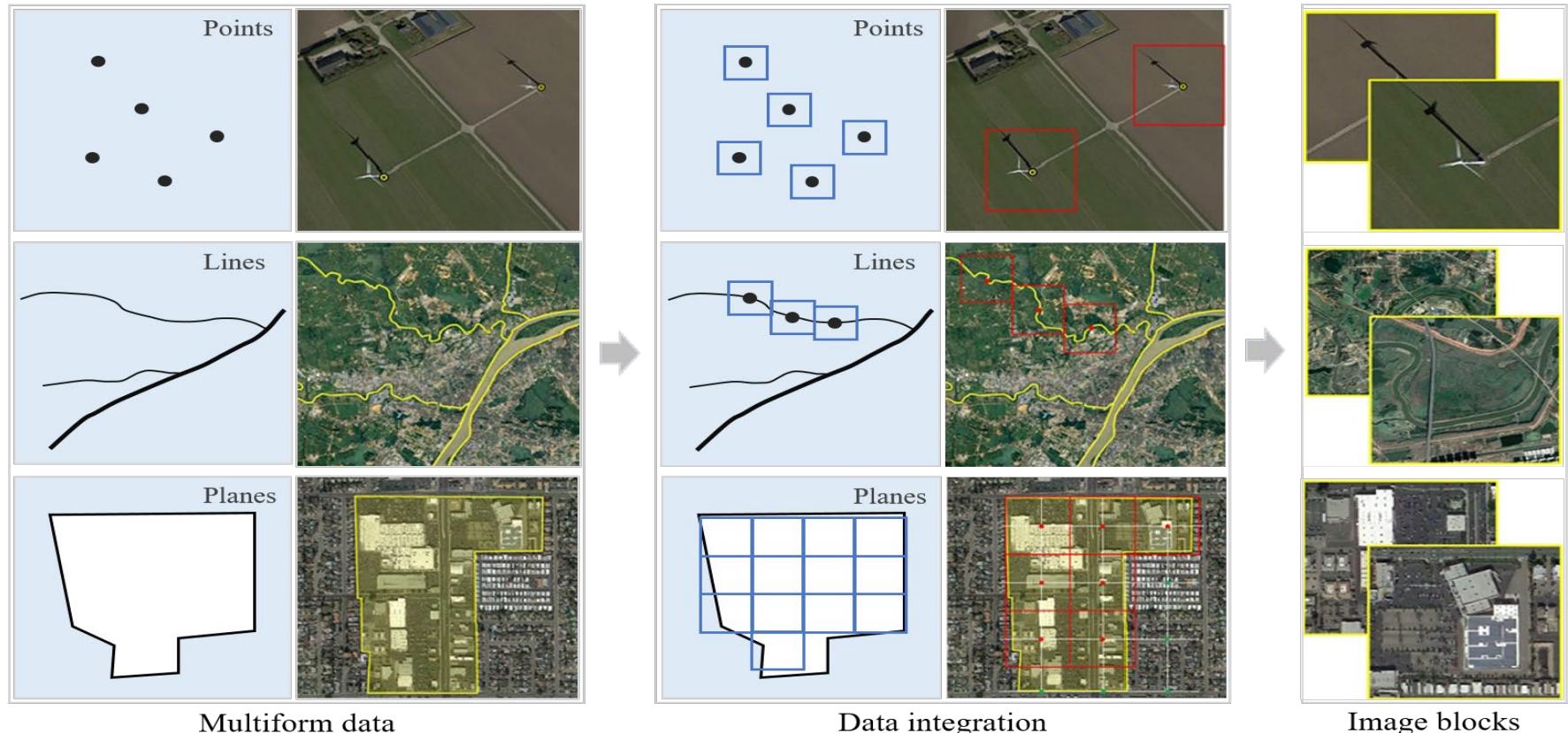
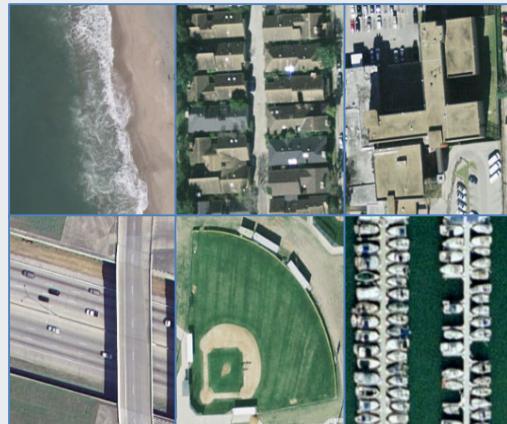


Illustration of the acquisition of RS scene images based on the collected geographic point, line and area data.
Points: centers of scene blocks. **Line:** sampled by intervals. **Plane:** sampled by mesh grids.

■ A Glimpse of Comparison

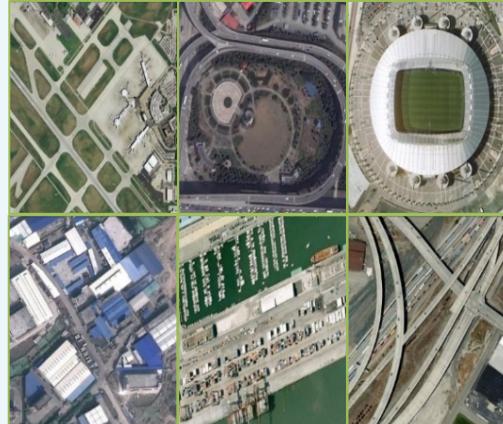
■ Million-AID: DiRS, better approximate real applications

UC-Merced



- Categories: 21
- Image size: 256x256
- Resolution: ~ 0.3m
- Number of images: 2100

WHU-RS19



- Categories: 19
- Image size: 600x600
- Resolution: 0.2 ~ 10m
- Number of images: 950

Million-AID



- Categories: 51
- Image size: 110~30,000
- Resolution: 0.2 ~ 153m
- Number of images: 1M



Outline

- Background
- Research Focus in the Past Decade
- Guidances to Benchmark RS Image Interpretation
- An Example: Million-AID
- Challenges and Perspectives
- Conclusions

- How to speed up the annotation process?

- **Visualization technology for RS image Annotation**

- *High-spectral images*: several selected bands, band transformation ...
- *Large-scale images*: efficient display technology helps to catch the content essence
- *SAR images*: signal expression via physical means

- **Annotation Efficiency and Quality Improvement**

- *Cooperation with application departments*: convert production data to algorithms
- *Annotation tools*: open-sourced and professional tools for RS image annotation
- *Noisy annotations*: noise cleansing, performance impact, tolerant algorithms

Challenges and Perspectives

■ Speed up the annotation process

Annotation tools for image dataset construction

No.	Name	Year	Description
1	LabelMe	2008	An online image annotation tool that supports various annotation primitives, including polygon, rectangle, circle, line and point.
2	Video Annotation Tool from Irvine, California (VATIC)	2012	An online tool that efficiently scaling up video annotation with crowdsourced marketplaces (e.g., AMT).
3	LabelImg	2015	A popular graphical image annotation application that labels objects in images with bounding boxes.
4	Visual Object Tagging Tool (VOTT)	2017	An open source annotation and labeling tool for image and video assets, extensible for importing/exporting data to local or cloud storage providers, including Azure Blob Storage and Bing Image Search.
5	Computer Vision Annotation Tool (CVAT)	2018	A universal data annotation approach for both individuals and teams, supporting large-scale semantic annotation for scene classification, object detection and image segmentation.
6	Image Tagger	2018	An open source online platform to create and manage image data and diverse labels (e.g., bounding box, polygon, line and point), with friendly support for collaborative image labeling.
7	Polygon RNN++	2018	A deep learning-based annotation strategy, producing polygonal annotation of objects segmentation interactively using humans-in-the-loop.
8	Makesence.AI	2019	An open source and online image annotation platform, using different artificial model to give recommendations as well as automate repetitive and tedious labeling activities.
9	VGG Image Annotator (VIA)	2019	A simple and standalone manual annotation software for image and video, providing rich labels like point, line, polygon as well as circle and ellipse without project management.



Outline

- Background
- Research Focus in the Past Decade
- Guidances to Benchmark RS Image Interpretation
- An Example: Million-AID
- Challenges and Perspectives
- **Conclusions**

■ A review of annotated datasets for RS image interpretation

- Covering literature published over the past decade
- A systematic review of the existing RS image datasets concerning the current mainstream of RS image interpretation tasks

■ Guidances to build RS image benchmarks

- DiRS: on creating benchmark datasets for RS image interpretation
- A picture of coordinates collection, methodology for RS image dataset construction

■ An example for dataset construction : Million-AID

- A large-scale benchmark dataset for RS image scene classification



CAPTAIN
COMPUTATIONAL AND PHOTOGAMMETRIC VISION

THANKS



*School of Computer Science, Wuhan University
Institute of Artificial Intelligence, Wuhan University
State Key Lab. LIESMARS, Wuhan University*

Gui-Song Xia (guisong.xia@whu.edu.cn)