frontiers

Supplementary Material

Deep Learning-based Semantic Segmentation of Remote Sensing Images: A Review

Jinna Lv, Qi Shen*, Mingzheng Lv, Yiran Li, Lei Shi, Peiying Zhang

*** Correspondence:** Qi Shen：sftshenqi@buu.edu.cn

1 Supplementary Tables

Table 1. Remote sensing datasets for semantic segmentation.

DataSets	Year	Descriptions	Channels	Sizes	Classes	Links to datasets or papers
ISPRS Vaihingen	2012	It contains many independent small buildings and provides 33 true ortho-photo (TOP) images in high resolution.	4-band(VINR)	2500×2000	5 classes: impervious surface, building, low vegetation, tree, car, and clutter.	https://www2.isprs.org/commissions/comm2/wg4/benchmark/semantic-labeling/
ISPRS Potsdam	2012	It consists of 38 image tiles, which cover an area of 3.42 km2 with a spatial resolution of 5 cm.	4-band(VINR)	6000×6000	5 classes: background, car, tree, low vegetation, building, impervious surfaces.	https://www2.isprs.org/commissions/comm2/wg4/benchmark/semantic-labeling/
GID	2018	It covers 506 km2 areas which are captured by the Gaofen-2 satellite. The dataset has a rich diversity in spectrum, texture, and structure.	GF-2	56×56	5 classes: buildings, farmland, forests, grasslands, and waters.	https://x-ytong.github.io/project/GID.html
WHDLD	2018	It is captured from Wuhan urban area.	Red/Green/Blue/IR	256×256	6 classes: building, road, pavement, vegetation, bare soil, and water.	https://sites.google.com/view/zhouwx/dataset#h.p_hQS2jYeaFpV0
iSAID	2020	It is the first benchmark dataset for instance segmentation in aerial images. 655,451 instances of the real object are collected in 15 classes across 2,806 images.	GF-2	800×800~13000×13000	15 classes: small vehicle, ground track field, swimming pool, bridge, etc.	https://captain-whu.github.io/iSAID/index.html
Massachusetts Road	2013	It covers 2600 km2 of Massachusetts.	Red/Green/Blue	1500×1500	2 classes: road and non-road.	https://www.cs.toronto.edu/~vmnih/data/
DLRSD	2022	A dense labeling dataset; Gt format; single channel 8bit color image; recommended to use PIL reading.	single channel 8-bit color image	256×256	17 classes: airplane, bare soil, building, car, chaparral, court, dock, field, grass, mobile home, pavement, etc.	https://sites.google.com/view/zhouwx/dataset#h.p_hQS2jYeaFpV0

RSSCN7 Data Set	2015	The dataset contains 2800 remote sensing images, which contains 400 images based on 4 Sampling at different scales.	Red/Green/Blue	400×400	7 classes: grassland, forest, farmland, parking lot, residential region, industrial region, and river/lake.	https://sites.google.com/site/qinzoucn/documents
LoveDA	2021	It offers 5987 high spatial resolution (0.3 m) remote sensing images from Nanjing, Changzhou, and Wuhan. Besides, there exist three considerable challenges: Multi-scale objects; Complex background samples; Inconsistent class distributions.	Red/Green/Blue	1024×1024	7 classes: building, road, water, infertile soil, forest, and agriculture.	https://github.com/Junjue-Wang/LoveDA
Landcover dataset	2020	It provides images of Poland's rural areas of 216.27 km2 in total. 39.51 km2 with 50 cm/pixel resolution and 176.76 km2 with 25 cm/pixel resolution are included.	Red/Green/Blue	9000 × 9500, 4200 × 4700	3 classes: buildings, forest land, and water.	https://landcover.ai/
DeepGlobe Land Cover Dataset	2018	It provides high-resolution sub-meter satellite imagery with a focus on rural areas.	Red/Green/Blue	2448 × 2448	7 classes: urban, agriculture, rangeland, forest, water, barren, unknown.	http://deepglobe.org/index.html
RIT-18	2017	Tetracam Micro-MCA6 multispectral imaging sensor flown onboard a DJI-1000 octocopter.	6-band(VNIR)	9393 × 5642, 8833 × 6918, 12446 ×7654	18 classes: asphalt, water pond, water lake, sand beach, grass laws, low-level vegetation, rocks, bunny, etc.	https://github.com/rmkemker/RIT-18
AIR-SEG	2020	57 labeled images collected from Google Earth are included in this dataset. It also offers 72 unlabeled images with a spatial resolution of 0.27 m.	Red/Green/Blue	2000×2000	10 classes: impervious surfaces; factory; residential building; road; water; farmland; bare land; grass; tree; clutter.	https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9185011
AISD	2022	It collects from Open Street Map online. It includes images from six districts: Chicago, Potsdam Paris, Zurich Berlin.	Red/Green/Blue/IR	3000 × 3000	3 classes: background, road, and building.	https://ui.adsabs.harvard.edu/abs/2022IJGI...11..165S/abstract
Beijing Land-Use Data Set	2018	It was collected by the Beijing-2 constellation in June 2018.	Red/Green/Blue	47244×47244	6 classes: construction area, tree, water, farmland, grass, bare soil.	https://ieeexplore.ieee.org/document/8970467/
EvLab-SS	2017	It has 60 images captured by different satellites with a resolution between 0.1 and 2m.	WorldView2/GeoEye	2000 × 1777 , 7210 × 5408	11 classes: background, farmland, garden, woodland, grassland, building, road, structures, digging pile, desert, and waters.	http://earthvisionlab.whu.edu.cn/zm/SemanticSegmentation/index.html

SIRI-WHU Dataset	2016	It consists of 31500 remote sensing images divided into 45 scene classes.	Red/Green/Blue	256 × 256	45 types of land use.	https://gcheng-nwpu.github.io/#Datasets
UAVid	2020	It focuses on city street scenes and contains 300 images intensively classified into eight classes to cope with the semantic labeling task.	Red/Green/Blue	4096 × 2160, 3840 × 2160	8 classes: building, road, trees, low vegetation, moving cars, static cars, human being, and clutter/background	https://www.uavid.nl/
DroneDeploy	2019	The dataset comprises numerous aerial scenes captured by drones. Every scene owns a ground resolution of 10 cm per pixel.	Red/Green/Blue	6000±× 6000±	7 classes: buildings, sundries, vegetation, water, ground, cars, etc.	https://github.com/dronedeploy/dd-ml-segmentation-benchmark
Barley Remote Sensing Dataset	2022	It displays a rural area in Xingren City, Guizhou Province, China, containing numerous crop fields. It was generated by an Unmanned Aerial Vehicle (UAV) near the ground.	Red/Green/Blue	47161 × 50141, 77470 × 40650	4 classes: flue-cured tobacco, corn, barley, and artificial building.	https://tianchi.aliyun.com/competition/entrance/231717/information
UCM	2018	It includes 2100 images with 0.3m spatial resolution.	Red/Green/Blue	256×256	17 classes: vegetation, ground, pavement, building, water, airplane, car, ship, etc.	https://pan.baidu.com/s/1mjPToHq#list/path=%2F
FloodNet	2021	The data is collected using the small UAV platform DJI Mavic Pro four-axis aircraft. There are 2343 images in the whole dataset.	Red/Green/Blue	3000×4000	10 classes: background, buildings are flooded, buildings are not flooded, roads are flooded, roads are not flooded, water, trees, vehicles, pools, grass.	https://github.com/BinaLab/FloodNet-Supervised_v1.0
GF-2 images	2020	This dataset comprises images of the Chenzhou area on the basis of the GF-2 satellite. It's an elaborate dataset that is well protected with intact primary secondary forest communities and broad-leaved forests at low-altitude in the Nanling Mountains.	Red/Green/Blue/IR	2000×2000	7 classes: forest, others, water, house, road, furrow, background.	https://directory.eoportal.org/web/eoportal/satellitemissions/g/gaofen-2
Flevoland dataset	1989	The Flevoland dataset is acquired by AIRSAR airborne platform in August 1989.	RGB	1024 × 750	15 classes: stem beans, peas, forest, Lucerne, three types of wheat, beet, potatoes, bare soil, grass, rapeseed, barley, water, and a small number of buildings.	https://download.csdn.net/download/qqinrui/11985726

Ottawa Dataset	1997	It consists of a pair of SAR images over the city of Ottawa of section pixels attained by the RADARSAT SAR sensor.	SAR	290×350	2 classes: water and background.	https://link.springer.com/article/10.1007/s11063-019-10174-x
Zeebrugge dataset	2015	This dataset derives from the 2015 IEEE GRSS Data Fusion Contest. The data scenario of this set belongs to the whole city and port of Zeebrugge.	IDs Tile 4 and Tile 6	10000×10000, 5000×5000	8 classes: Impervious Surface, buildings, low vegetation, tree, car, clutter, boat, and water.	http://www.grss-ieee.org/community/technical-committees/data-fusion
WHU Aerial Building	2011	It comprises 22,000 independent buildings. Besides, the ground resolution of the images is 0.075m.	0.075 m	512×512	2 classes: building and background.	https://data.linz.govt.nz/layer/51932-christchurch-post-earthquake-01m-urban-aerial-photos-24-february-2011/
Lake water dataset	2022	The dataset after data augmentation processing contains 2733 remote sensing images with a size of 416 ×416.	Red/Green/Blue	416 ×416	2 classes: lake and not lake.	https://link.springer.com/article/10.1007/s10489-022-03345-2?utm_source=xmol&utm_content=meta
UAV-borne dataset	2022	With a resolution between 10000×10000 and 20000×20000 is constructed. It's from a city in Guangdong and a place along the Yangtze River.	Red/Green/Blue	512×512	2 classes: road and building.	https://readpaper.com/paper/4702452835208413185
Pavia datasets	2021	It was acquired by the ROSIS sensor over the campus of Pavia University and its surrounding city area, Pavia, Italy.	Red/Green/Blue	610×340, 1096×715	8 classes: asphalt meadows, tiles, trees, water, bare soil, bitumen, bricks, and shadows.	https://rslab.ut.ac.ir/data
Local Climate Zone DataSets	2017	It is composed of six large-scale city scenes: Munich, Berlin, Hong Kong, Paris, Rome, and Sao Paulo.	10 m and ten channels	7064×6528	8 classes: compact mid-rise, open mid-rise, open low-rise, large low-rise, dense trees, scattered trees, low plants, and water.	http://www.grss-ieee.org/2017-ieee-grss-data-fusion-contest/
Quickbird image	2020	The QuickBird images are in Hangzhou City and Nanjing City, China.	Red/Green/Blue and near-infrared bands.	58×504, 538×546, 996×550	2 classes: boundary and non-boundary.	https://www.sciencedirect.com/science/article/abs/pii/S0924271619302631

South Korea Building Dataset	2021	The regions covered by the dataset are located in Ansan and Siheung, Korea, where houses, apartments, and factories all exist.	Red/Green/Blue	9122 × 8892	2 classes: buildings and non-buildings.	https://www.mdpi.com/2072-4292/13/16/3087
Paris-Lille-3D	2018	It is a Dataset and a Benchmark on Point Cloud Classification. The data has been produced by a Mobile Laser System (MLS) in two different cities in France (Paris and Lille).	3D point cloud	-	50 classes: natural, cars, pedestrians, barriers, trash cans, bollards, poles, buildings, ground, etc.	http://caor-mines-paristech.fr/fr/paris-lille-3d-dataset
Namtso Lake data	2020	Satellite data and application of the Qinghai-Tibet Plateau and a river in central China from 2015 to 2019 by China Resources Center.	Red/Green/Blue	512 × 512	2 classes: Lake and background.	http://www.cresda.com/CN/

Table 2. Comparison of different methods on ISPRS Potsdam and Vaihingen datasets.

Methods	Models	ISPRS Potsdam			ISPRS Vaihingen		
		mF1(%)	mIoU(%)	OA(%)	mF1(%)	mIoU(%)	OA(%)
Base on CNN	EFCNet[68]	79.74	65.7	80.72	81.87	70.14	85.46
	EGCAN[78]	93	-	91.4	89.7	-	90.1
	HCANet[56]	88.07	-	88.92	88.94	-	89.71
Attention Mechanism	MANet[29]	92.9	86.95	91.32	90.41	82.71	90.96
	ABCNet[100]	92.7	86.5	91.3			
	A2-FPN[101]	92.4	86.1	91.1	90.1	82.2	91
	LANet[109]	91.95	-	90.84	88.09	-	89.83
	SCAttNet[113]	-	68.31	87.97	-	-	-
	CAM-DFCN[106]	89.43	-	90.26	88.55	-	90.41
	DSPCANet[31]	-	77.66	90.13	-	72.56	87.32
	MAResUNet[28]	-	-	-	90.28	83.3	90.86
Multiscale Strategy	FSHRNet[24]	90.67	83.16	89.82	86.66	88.38	76.86
	SFAMNet[26]	88.18	71.31	86.77	86.91	65.28	88.45
	MFNet[27]	-	-	91.65	88.24	77.05	91.47
	DGPRNet[89]	-	77.05	85.69	-	82.36	90.43
Based on Transformer	DC-Swin[32]	93.25	87.56	92	90.71	83.22	91.63
	DHT-E[140]	-	81.7	89.3	-	-	-
	ICTNet[135]	93	-	91.57	92.34	-	90.14
	MAT [142]	91.59	84.82	-	88.7	79.93	-
	SUDNet[137]	92.57	86.4	92.98	89.49	81.26	90.95
	CG-Swin[132]	93.29	87.61	91.93	90.81	83.39	91.68
	SwinTF-PSP[134]	-	-	-	94.83	90.98	-
Based on GAN	Semi-supervised [145]	88.57	-	87.89	87.08	-	88.34