

Automated Detection of Antarctic Benthic Organisms in High-Resolution In Situ Imagery to Aid Biodiversity Monitoring

Supplementary Material

A. Dataset Composition

Table S1. Counts and areas for the Weddell Sea Benthic Dataset classes, for both the original large-scale and patched images. Bolded counts denote the most abundant classes used for 10-class evaluation.

Class Label	Whole Image Dataset							Patched Dataset						
	Count				Area (px ²)			Count				Area (px ²)		
	All	Train	Validation	Test	Min	Max	Avg	All	Train	Validation	Test	Min	Max	Avg
actinarian	165	118	22	25	111	51328	6821	777	553	105	119	111	51328	5512
alcyonium	280	266	6	8	266	104355	6773	1269	1202	24	43	60	98203	5508
anthomastus	89	60	22	7	198	30478	4666	371	260	80	31	106	30478	3810
ascidian.cnemidocarpa.verrucosa	10	2	4	4	4689	91344	21113	55	15	20	20	1517	91344	15311
ascidian.distaplia	32	24	1	7	1303	1432444	68092	186	146	4	36	539	250000	34927
ascidian.pyura.bouvetensis	66	41	21	4	981	188218	16652	351	203	126	22	275	153083	12022
asteroidia	156	111	23	22	144	569242	9572	643	482	74	87	26	250000	8279
astrochlamys	720	528	127	65	475	136010	13402	3366	2439	556	371	110	126777	10608
benthic_fish	71	52	14	5	309	179118	40603	474	349	92	33	309	139500	23828
bryozoan	15	8	5	2	386	45004	14134	88	45	35	8	386	45004	9918
crinoid	26	19	3	4	491	20989	6422	119	78	14	27	232	20989	5377
crustaceans	461	338	79	44	554	399481	12799	2469	1817	424	228	253	173560	9494
cucumber	355	287	41	27	62	9041	1522	1447	1166	164	117	56	9041	1415
cup_coral	4757	2807	1611	339	31	13756	520	18552	11262	6039	1251	12	13756	482
demosponges	2211	1517	340	354	7	258424	2283	8960	6216	1312	1432	7	194988	2003
echinoid	11	6	2	3	1975	12170	4098	50	24	12	14	817	12170	3607
glass.sponge	2308	1612	477	219	29	92647	1603	9144	6295	1930	919	13	92647	1508
gorgonian	1144	903	113	128	62	303363	9045	5396	4248	546	602	62	219445	7274
hydroid_solitary	25	17	5	3	2519	97165	14968	133	94	25	14	378	97165	11492
ophiosabine	3075	1853	694	528	219	50859	3125	12783	7630	2897	2256	68	50859	2768
ophiuroid_5_arms	1885	1293	280	312	93	748890	17419	8819	5961	1349	1509	72	250000	13686
pencil_urchin	78	51	20	7	700	36549	6419	338	219	86	33	130	36549	5296
pycnogonid	11	7	2	2	2381	25210	13854	59	40	9	10	747	25210	9264
stylasterids	13295	9547	2011	1736	4	74999	1720	53523	38529	7543	7451	4	69524	1560
worm_tubes	35	19	2	14	168	25355	4591	157	82	8	67	168	25355	3958
Total	31280	21486	5925	3869				129529	89355	23474	16700			

B. Data Augmentation Strategy Details

Tab. S2 presents the Albumentations-based [6] data augmentation techniques used in this study, categorised by the augmentation strategy in which they were employed. Probabilities for all augmentations were set to 0.50. For *Random Sized BBox Safe Crop*, the height and width parameters were set to the patch size. All other parameters were set to the Albumentations default.

Despite its name, *Pixel Dropout* is classified as a spatial transformation. This augmentation operates non-uniformly across the image by randomly selecting specific spatial coordinates at which to drop pixels. Consequently, it alters the spatial structure of the image rather than applying a uniform change across all pixels.

Table S2. A list of data augmentation techniques provided by the Albumentations library, along with the augmentation strategy in which each technique was applied.

Augmentation	Strategy		
	Pixel	Spatial	Both
Horizontal Flip	✗	✓	✓
Motion Blur	✓	✗	✓
Pixel Dropout	✗	✓	✓
Random Brightness and Contrast	✓	✗	✓
Random Shadow	✓	✗	✓
Random Sized BBox Safe Crop	✗	✓	✓
Vertical Flip	✗	✓	✓

C. Weddell Sea Benthic Dataset High-Resolution Examples

Example Weddell Sea Benthic Dataset test set image outputs. Predicted organism bounding boxes, class labels, and confidence scores shown after reprojection and postprocessing. Confidence threshold = 0.60.

Figure S1. HOTKEY_2019_03_31_at_13_30_13_IMG.0853. Original size: 3799×3798 px.

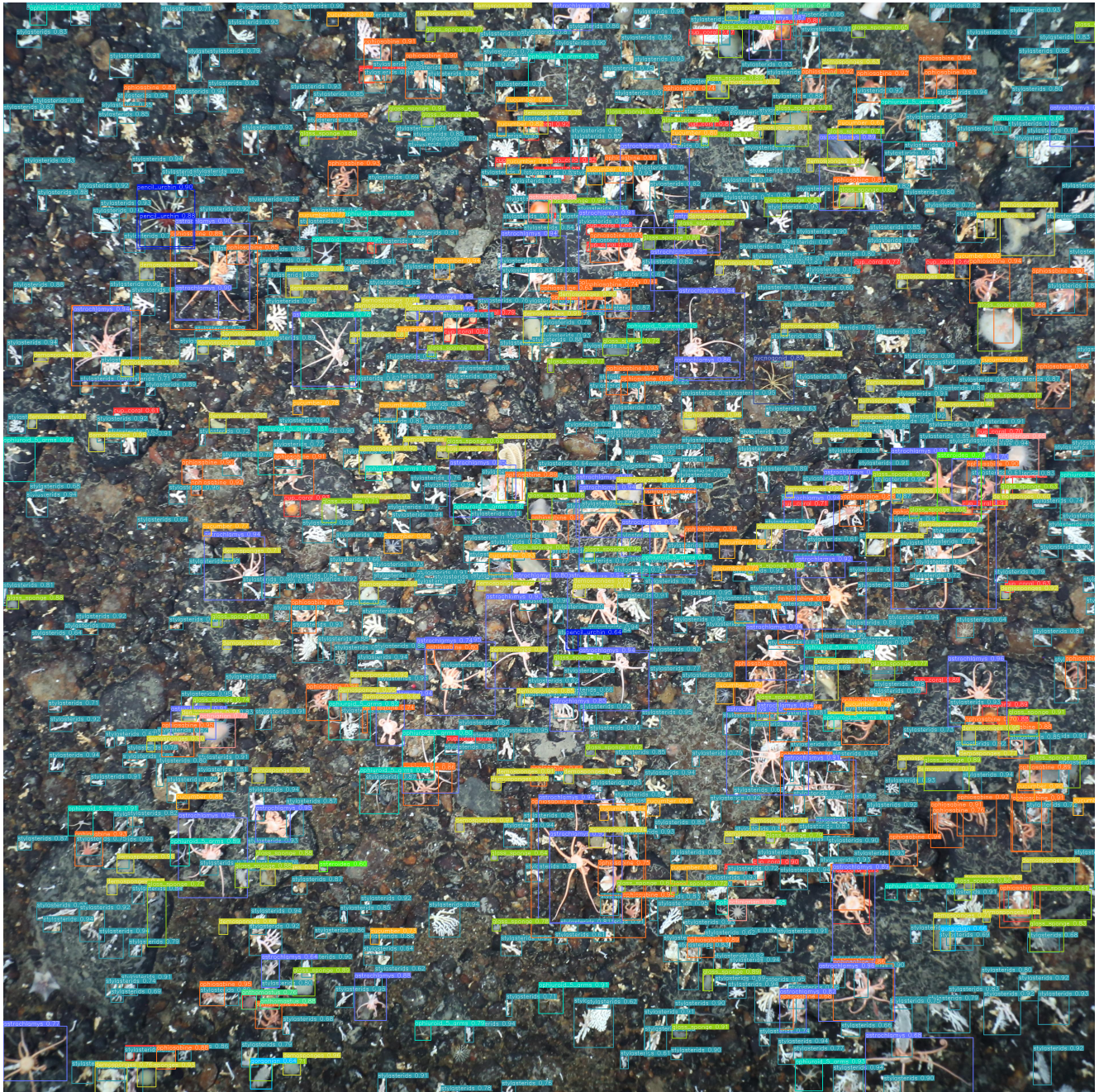
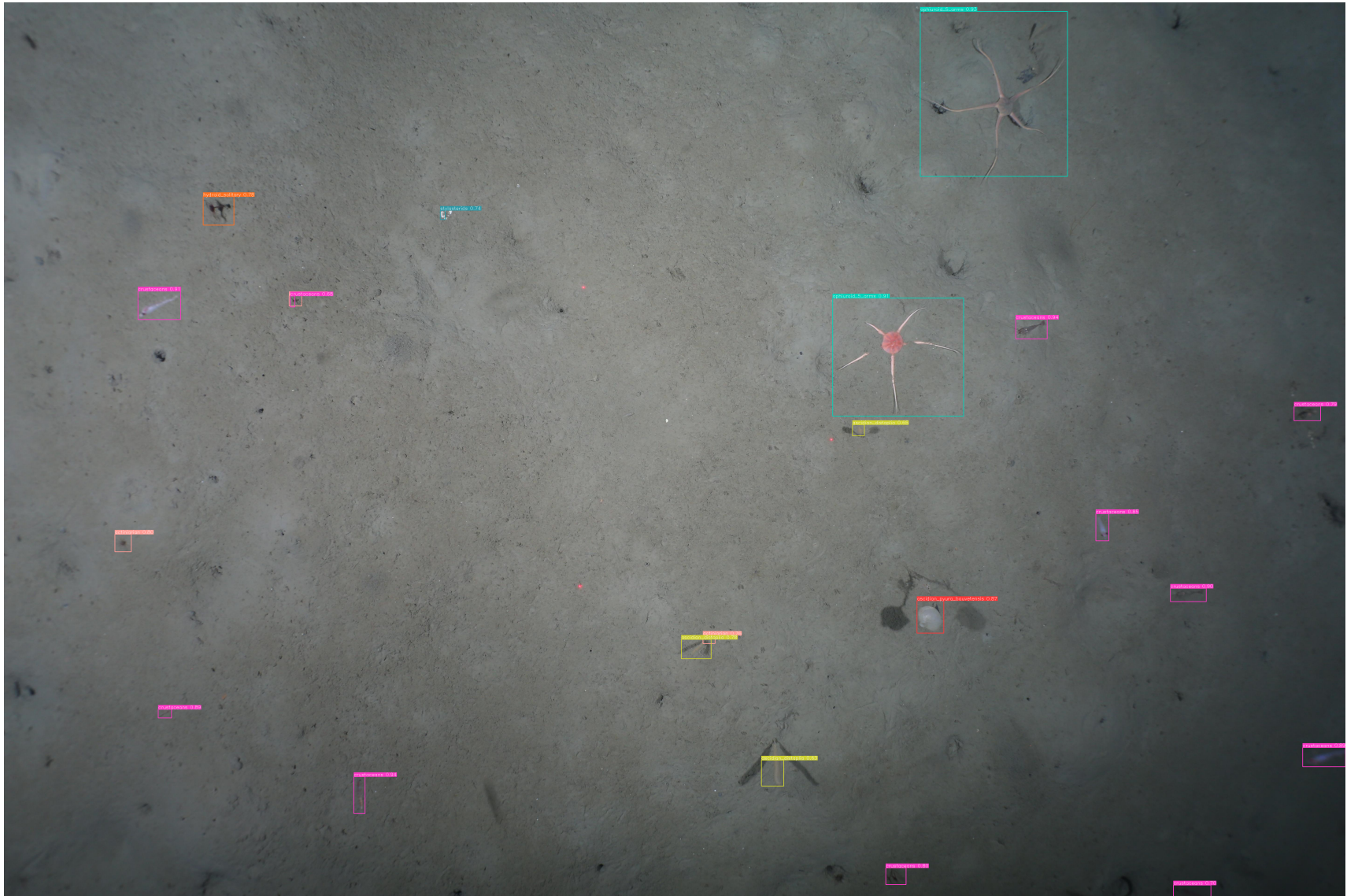


Figure S2. HOTKEY_2019_03_31_at_13_21_23_IMG.0816. Original size: 2975x2964 px.



Figure S3. TIMER_2019_03_06_at_05_40_47_IMG.0253. Original size: 5760×3840 px.



Organism abundance was shown to have a large effect on framework performance. See 6.2 for further discussion.

Abundance

	stylasterids	cup_coral	ophiosabine	glass_sponge	ophiuroid_5_arms	demosponges	gorgonian	astrochlamys	crustaceans	alcyonium	cucumber	asteroidea	benthic_fish	pencil_urchin	anthomastus	actinarian	ascidian_pyura_bouvetensis	ascidian_distaplia	hydroid_solitary	worm_tubes	crinoid	pycnogonid	bryozoan	echinoid	ascidian_cnemidocarpa_verrucosa	background
stylasterids	51%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	46%	
cup_coral	4%	38%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	54%	
ophiosabine	0%	71%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	25%	
glass_sponge	1%	0%	35%	0%	13%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	47%	
ophiuroid_5_arms	0%	0%	3%	0%	72%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	21%	
demosponges	0%	0%	5%	0%	38%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	54%	
gorgonian	7%	0%	2%	0%	0%	31%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	56%	
astrochlamys	4%	0%	10%	0%	4%	0%	55%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	24%	
crustaceans	0%	0%	0%	0%	0%	0%	86%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	11%	
alcyonium	0%	0%	0%	11%	0%	0%	0%	77%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	11%	
cucumber	2%	0%	8%	5%	2%	0%	0%	0%	41%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	38%	
asteroidea	13%	0%	6%	3%	3%	0%	0%	0%	30%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	43%	
benthic_fish	0%	0%	16%	0%	0%	0%	0%	0%	0%	0%	50%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	33%	
pencil_urchin	0%	0%	0%	0%	9%	0%	0%	0%	0%	0%	27%	0%	0%	0%	0%	0%	0%	0%	27%	0%	0%	0%	0%	0%	36%	
anthomastus	0%	0%	11%	0%	0%	0%	0%	0%	0%	0%	0%	33%	0%	0%	0%	0%	0%	0%	0%	0%	0%	11%	0%	0%	44%	
actinarian	0%	5%	0%	0%	0%	0%	0%	0%	17%	0%	0%	0%	2%	29%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	44%	
ascidian_pyura_bouvetensis	0%	0%	20%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	40%	0%	0%	0%	0%	0%	0%	0%	0%	0%	40%	
ascidian_distaplia	0%	0%	0%	22%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	11%	0%	0%	0%	0%	0%	0%	0%	0%	66%	
hydroid_solitary	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	25%	0%	0%	25%	0%	0%	0%	0%	0%	0%	0%	0%	50%	
worm_tubes	0%	0%	0%	0%	6%	0%																				

E. SAHI Postprocessing Limitations

Large objects, split over a high number of patches, may fail to merge into a single coherent bounding box after Non-Maximum Merging via SAHI. See 6.3 for further discussion. Confidence threshold = 0.60.

Figure S5. A single ophiuroid_5_arms, represented by two bounding boxes after postprocessing. Cropped and enlarged for clarity.

