Table S1. Predefined morphological classes and their corresponding mask value, asterisks indicate 17 classes that were predicted during training.

Class ID	Label
1	**Live_Branching***
2	Live_Digitate
3	Live_Columnar
4	Live_Enrusting_Columnar
5	**Live_Encrusting***
6	**Live_Porites***
7	Live_FreeLiving
8	**Live_Foliose***
9	**Live_Massive***
10	**Live_SubMassive***
11	**Live_Tabulate***
12	**Dead_Branching***
13	**Dead_Foliose***
14	**Dead_Massive***
15	**Dead_Tabulate***
16	NonCoral_Branching
17	NonCoral_Encrusting
18	NonCoral_Massive
19	NonCoral_SubMassive
20	**Rubble***
21	**SoftCoral_Nodular***
22	SoftCoral_Plating
23	**Macroalgae***
24	Dead_Columnar
25	**Sand***
26	**Dead_SoftCoral***
27	**Live_Massive_Grooved***

Table S2. Number of images and tiles and corresponding training, validation, and testing tiles used for model development across survey years. Data from 2015, 2017, and 2019 were split into training, validation, and testing sets, while 2023 data were reserved exclusively for out-of-sample evaluation. The final row shows the total number of tiles used across all years.

Year	# of images	Training tiles	Validation tiles	Testing tiles	Total
2015	24	27,468	5,886	5,886	39,240
2017	29	33,190	7,112	7,112	47,415
2019	24	27,468	5,886	5,886	39,240
Total	**77**	**88,126**	**18,884**	**18,884**	**125,901**
2023	27	n/a	n/a	n/a	44,145
Total	104				170,049

Table S3. Class-level performance metrics for the CNN-based benthic classification model, computed across all KI15, KI17, and KI19 plots. Metrics include precision, recall, F1-score, Intersection over Union (IoU), and total support (number of ground truth pixels) per class. A final row presents the overall weighted average across all valid classes.

Class ID	Label	Precision	Category	Recall	f1	IoU	Support
1	**Live_Branching***	**0.74**	**Living**	**0.95**	**0.83**	**0.72**	**70086**
2	Live_Digitate	0.00	Living	0.00	0.00	0.00	298
3	Live_Columnar	0.00	Living	0.00	0.00	0.00	648
4	Live_Enrusting_Columnar	0.00	Living	0.00	0.00	0.00	1193
5	**Live_Encrusting***	**0.89**	**Living**	**0.89**	**0.89**	**0.80**	**116680**
6	**Live_Porites***	**0.96**	**Living**	**0.95**	**0.95**	**0.91**	**199925**
7	Live_FreeLiving	0.00	Living	0.00	0.00	0.00	8588
8	**Live_Foliose***	**0.93**	**Living**	**0.97**	**0.95**	**0.91**	**125104**
9	**Live_Massive***	**0.90**	**Living**	**0.92**	**0.91**	**0.84**	**111168**
10	**Live_SubMassive***	**0.88**	**Living**	**0.89**	**0.88**	**0.79**	**62023**
11	**Live_Tabulate***	**0.95**	**Living**	**0.90**	**0.92**	**0.86**	**28936**
12	**Dead_Branching***	**0.95**	**Non-living**	**0.89**	**0.92**	**0.85**	**211686**
13	**Dead_Foliose***	**0.48**	**Non-living**	**0.93**	**0.63**	**0.46**	**60486**
14	**Dead_Massive***	**0.94**	**Non-living**	**0.92**	**0.93**	**0.87**	**1452725**
15	Dead_Tabulate	0.00	Non-living	0.00	0.00	0.00	52532
16	NonCoral_Branching	0.00	Living	0.00	0.00	0.00	29
17	NonCoral_Encrusting	0.00	Living	0.00	0.00	0.00	676
18	NonCoral_Massive	0.00	Living	0.00	0.00	0.00	421
19	NonCoral_SubMassive	0.35	Living	0.97	0.51	0.34	10564
20	**Rubble***	**0.90**	**Non-living**	**0.94**	**0.92**	**0.85**	**945536**
21	**SoftCoral_Nodular***	**0.07**	**Living**	**0.07**	**0.07**	**0.04**	**19408**
22	SoftCoral_Plating	0.00	Living	0.00	0.00	0.00	18391
23	**Macroalgae***	**0.95**	**Living**	**0.92**	**0.94**	**0.88**	**288778**
24	Dead_Columnar	0.00	Living	0.00	0.00	0.00	894
25	**Sand***	**0.88**	**Non-living**	**0.87**	**0.87**	**0.78**	**116984**
26	**Dead_SoftCoral***	**0.00**	**Non-living**	**0.00**	**0.00**	**0.00**	**6434**
27	**Live_Massive_Grooved***	**0.59**	**Living**	**0.77**	**0.67**	**0.50**	**25721**
Total	**All (only valid classes)**	**0.89**		**0.90**	**0.89**	**0.83**	**3935914**

Table S4. Mean classification performance metrics (precision, recall, F1-score, IoU, and overall accuracy) for each reef site, aggregated across all available years and image subsets (MPQs).

Site	Precision	Recall	f1	Support	IoU	Accuracy
Site 5	0.41	0.45	0.43	477,112	0.74	0.88
Site 8	0.42	0.45	0.44	599,482	0.76	0.92
Site 15	0.42	0.46	0.42	324,406	0.72	0.84
Site 19	0.45	0.49	0.46	241,934	0.76	0.86
Site 27	0.35	0.38	0.35	402,426	0.80	0.96
Site 30	0.41	0.43	0.42	417,160	0.77	0.92
Site 32	0.31	0.39	0.33	347,739	0.55	0.91
Site 34	0.40	0.39	0.39	463,180	0.64	0.83
Site 35	0.42	0.42	0.42	369,520	0.79	0.92
Site 37	0.38	0.45	0.40	292,955	0.69	0.95

Table S5. Per-class proportional differences between predicted and ground-truth benthic cover across survey years (2015, 2017, 2019). Values represent (Predicted – True) proportions, with positive values indicating overestimation by the model and negative values indicating underestimation.

Class ID	Class	2015	2017	2019
1	Live_Branching	0.0046	0.0088	0.0023
2	Live_Digitate	-0.00017	-8.46E-06	-5.31E-05
3	Live_Columnar	-0.00048	-1.69E-05	N/A
4	Live_Enrusting_Columnar	-0.00079	N/A	-0.00013
5	Live_Encrusting	0.0027	0.00013	-0.0017
6	Live_Porites	0.0047	0.00012	-0.0044
7	Live_FreeLiving	-0.0025	-0.0016	-0.0024
8	Live_Foliose	0.0080	0.00012	3.67E-05
9	Live_Massive	0.0045	0.00014	-0.0016
10	Live_SubMassive	0.0012	0.00028	-0.00075
11	Live_Tabulate	-0.00019	3.97E-05	-2.30E-05
12	Dead_Branching	-0.0029	-0.00061	-0.0080
13	Dead_Foliose	0.0027	0.035	0.00079
14	Dead_Massive	0.0041	-0.010	-0.00021
15	Dead_Tabulate	-0.0034	-0.031	-0.00019
20	Rubble	-0.0077	-0.0024	0.0263
21	SoftCoral_Nodular	-0.00041	0.00018	-0.00024
22	SoftCoral_Plating	-0.014	-3.19E-05	-0.00053
23	Macroalgae	-0.00061	0.00031	-0.0070
24	Dead_Columnar	-0.00061	0	-7.47E-05
25	Sand	-0.0015	0.0018	-0.0024
26	Dead_SoftCoral	-2.98E-05	-0.00291	-0.0016
27	Live_Massive_Grooved	0.002517	0.00189	0.001786

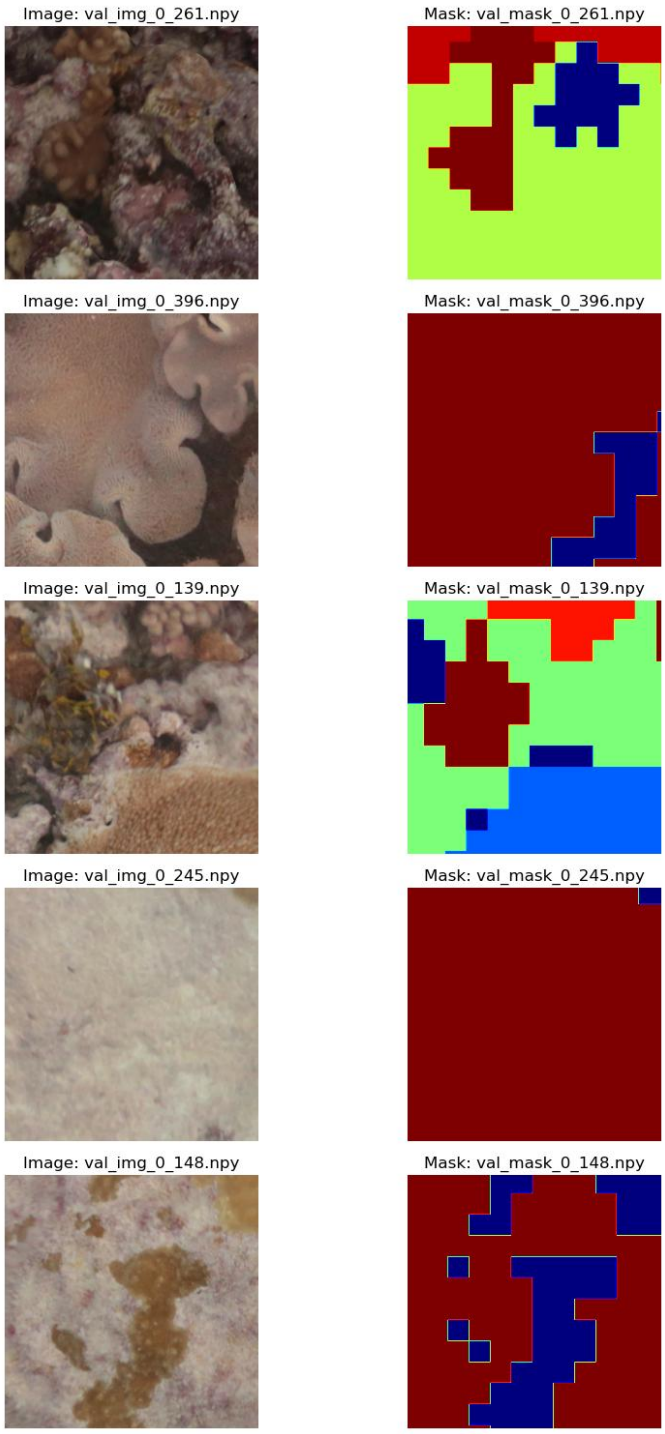


Figure S1. Examples of image tiles and their corresponding one-hot encoded outputs, demonstrating the 4 input format used for training the classifier.

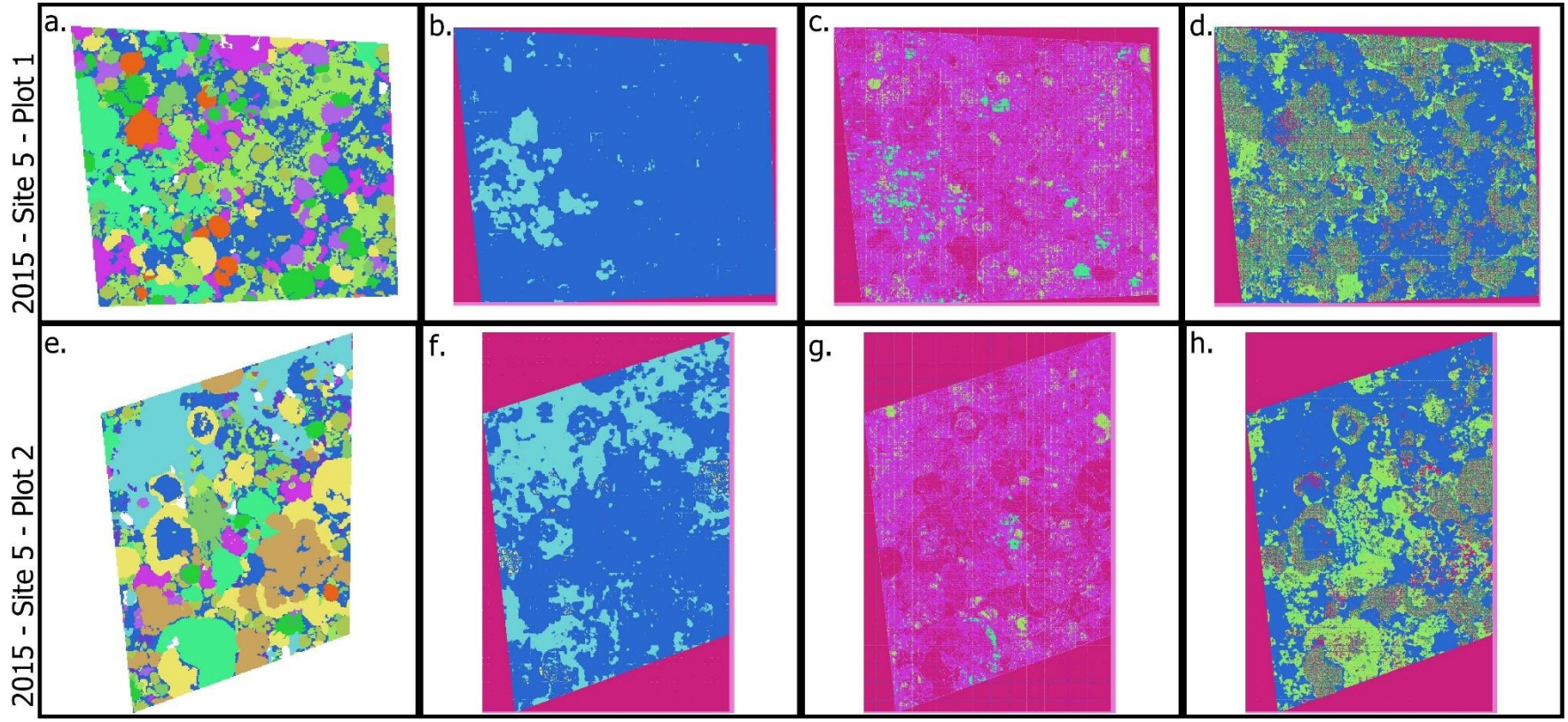


Figure S2. Examples from 2015 Site 5 plots (MPQs) 1 and 2 of models that performed poorly 6 under different loss functions, illustrating both overfitting and underfitting. Panels (a) and (e) show the 7 input mask and human annotation. Panels (b) and (f) show results trained using standard cross-entropy 8 loss with a restricted set of classes. Panels (c) and (g) show results using cross-entropy loss with class 9 weighting. Panels (d) and (h) show results from models trained with focal and Dice loss using weighted 10 classes.

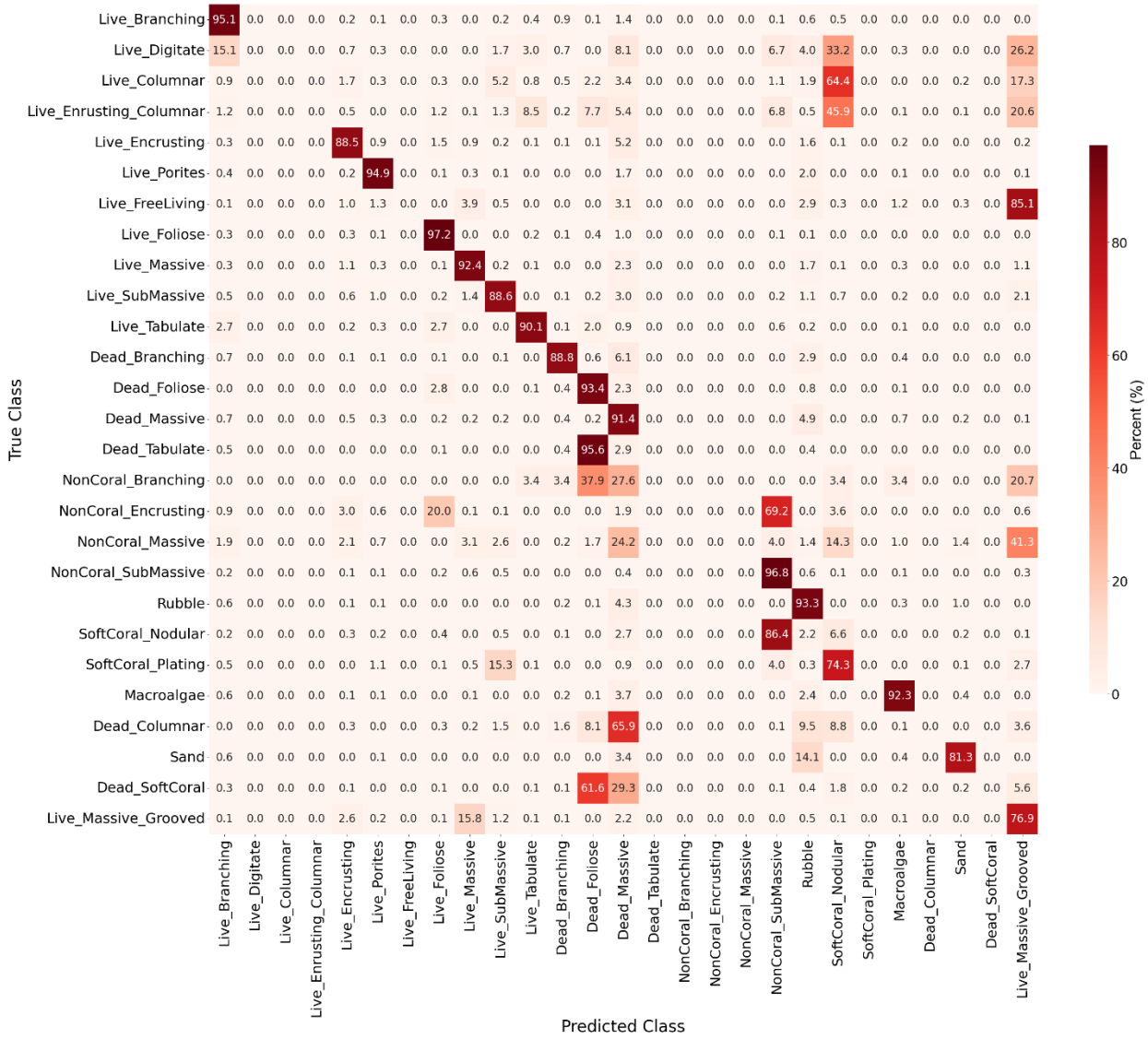


Figure S3. Normalized confusion matrix showing class-wise prediction accuracy as percentages. Each cell represents the proportion of pixels from a true benthic class (rows) that were predicted as each class (columns).

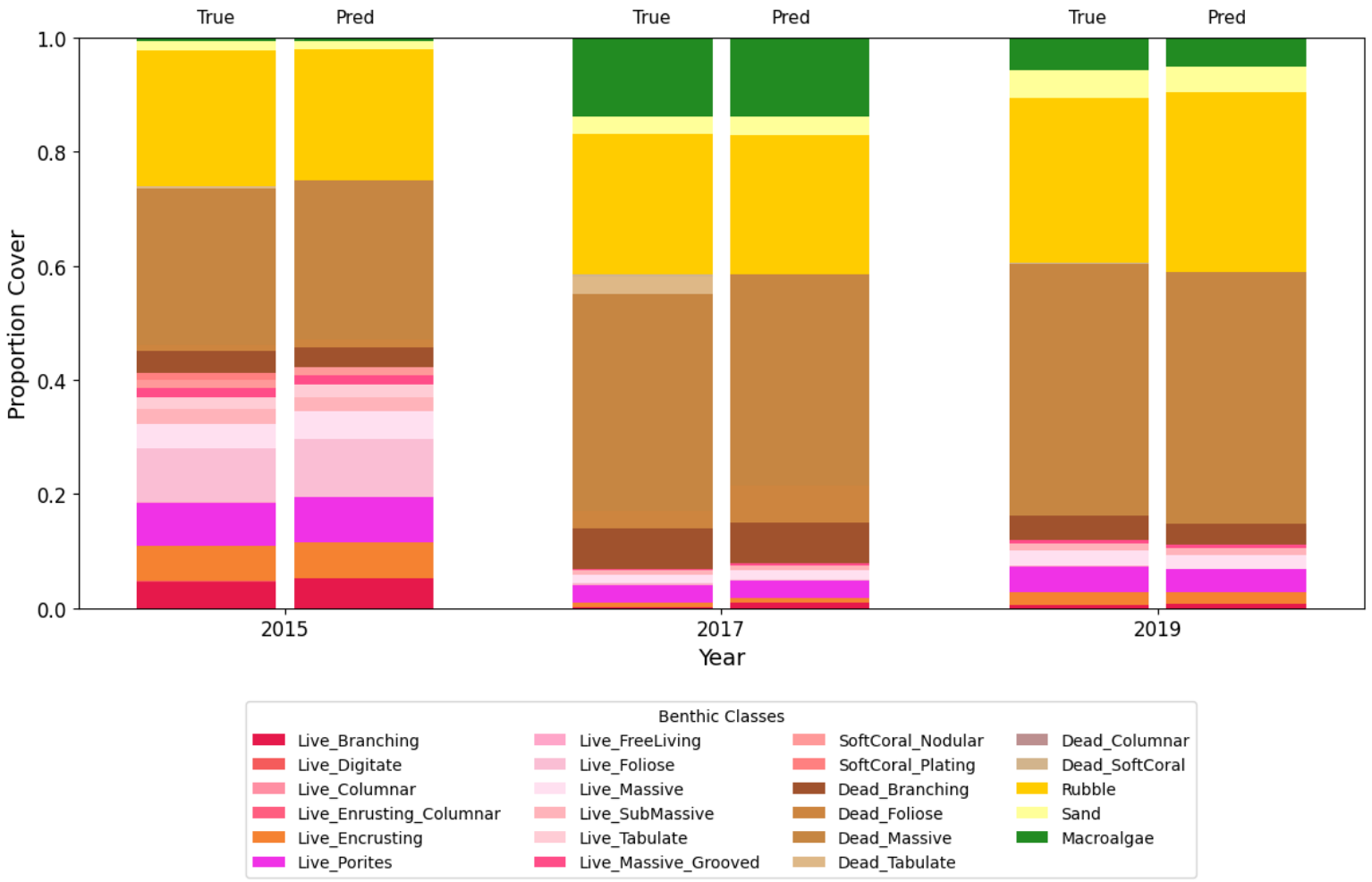


Figure S4. Stacked bar plots of benthic class proportions comparing ground truth and predicted masks across survey years 2015 (KI15), 2017 (KI17), and 2019 (KI19). For each year, class-specific cover is shown as proportions of total annotated pixels, with True values (left bars) derived from manual annotations and Pred values (right bars) from model predictions. Colors correspond to individual benthic classes.