Strategies to Improve Real-World Applicability of Laparoscopic Anatomy Segmentation Models

Fiona R. Kolbinger
Purdue University
fkolbing@purdue.edu

Jiangpeng He
Purdue University
he416@purdue.edu

Jinge Ma
Purdue University
ma859@purdue.edu

Fengqing Zhu
Purdue University
zhu0@purdue.edu

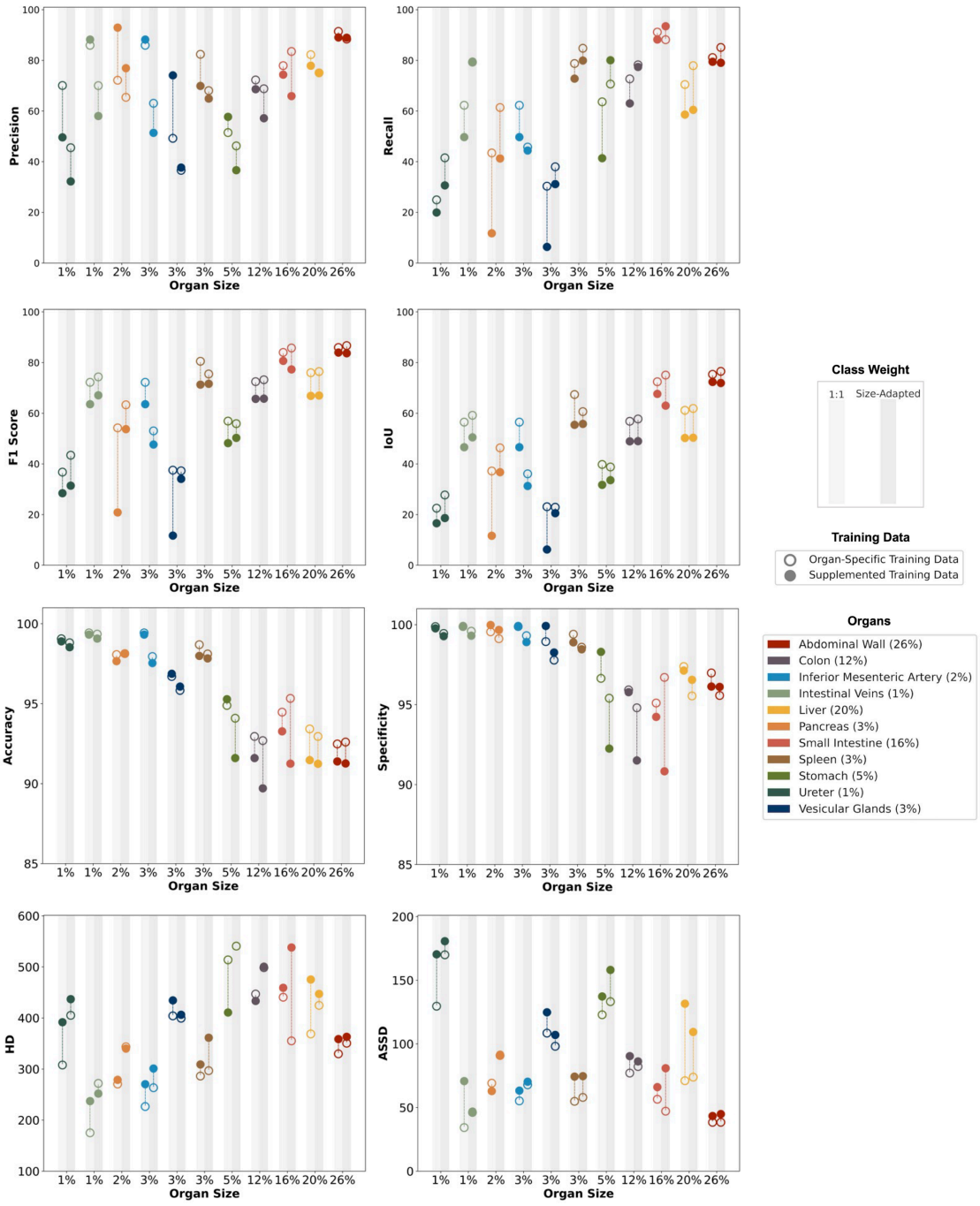
Supplementary Material

Supplementary Figure 1: Quantitative evaluation of segmentation model performance in relation to training data composition and foreground-background class weight, tested on organ-specific test data.. 2

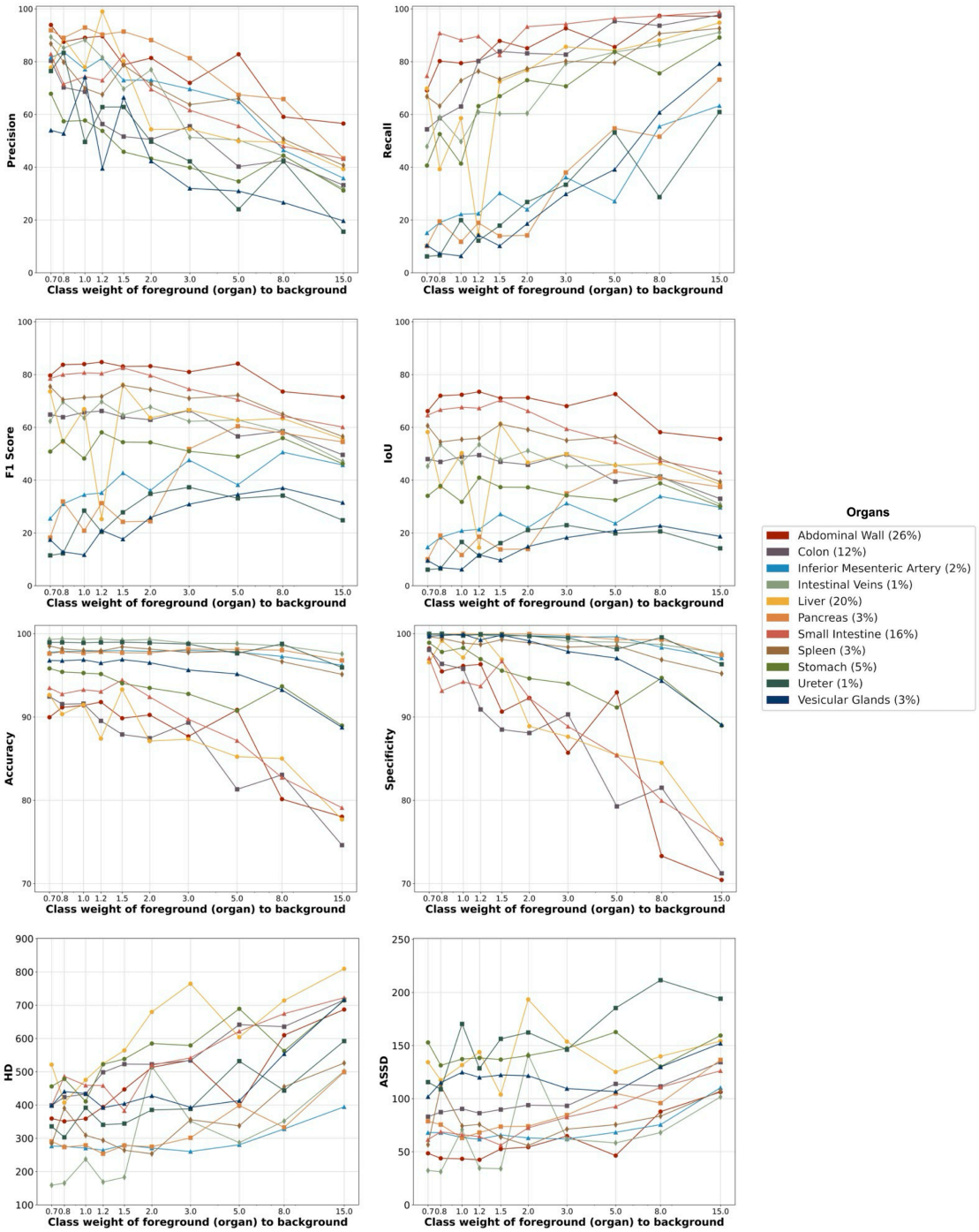
Supplementary Figure 2: Impact of variations in foreground-background class weights on segmentation performance metrics, tested on organ-specific test data..... 3

Supplementary Table 1: Performance of binary organ segmentation models trained with class weights identical for foreground and background and tested on organ-specific test data. 4

Supplementary Table 2: Performance of binary organ segmentation models trained with class weights adapted to organ size and tested on organ-specific test data 4



Supplementary Figure 1: Quantitative evaluation of segmentation model performance in relation to training data composition and foreground-background class weight, tested on organ-specific test data. Models were trained on either organ-specific training data including only class-positive images or supplemented training data including class-positive and class-negative examples. Foreground and background were either weighted identically (1:1) or adapted to organ size. Models were tested on organ-specific test data. Abbreviations: Average Symmetric Surface Distance (ASSD), Hausdorff Distance (HD), Intersection-over-Union (IoU).



Supplementary Figure 2: Impact of variations in foreground-background class weights on segmentation performance metrics, tested on organ-specific test data. Models were trained on supplemented training data including class-positive and class-negative examples and tested on organ-specific test data. Abbreviations: Average Symmetric Surface Distance (ASSD), Hausdorff Distance (HD), Intersection-over-Union (IoU).

Supplementary Table 1: Performance of binary organ segmentation models trained with class weights identical for foreground and background and tested on organ-specific test data. Models were trained on organ-specific (O) training data including only class-positive images or supplemented (S) training data including class-positive and class-negative examples. All models were tested on organ-specific test data. Abbreviations: Average Symmetric Surface Distance (ASSD), Hausdorff Distance (HD), Intersection-over-Union (IoU), Organ-specific training dataset (O), Supplemented training dataset (S).

Organ	Organ Size (% Foreground)	Accuracy		Precision		Recall		IoU		F1 Score		Specificity		HD		ASSD	
		O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S
Ureter	1.2%	99.06	98.90	70.07	49.60	24.92	19.94	22.52	16.58	36.76	28.45	99.88	99.77	307.90	391.73	129.60	170.35
Intestinal Veins	1.3%	99.43	99.32	85.94	88.20	62.26	49.69	56.50	46.60	72.20	63.57	99.88	99.92	175.05	237.23	34.26	70.78
Pancreas	2.7%	98.07	97.66	72.15	92.92	43.44	11.74	37.20	11.64	54.23	20.85	99.55	99.98	270.60	278.81	69.12	62.91
Inferior Mesenteric Artery	2.8%	99.43	99.32	85.94	88.20	62.26	49.69	56.50	46.60	72.20	63.57	99.88	99.92	226.56	270.4	55.21	63.27
Vesicular Glands	2.9%	96.70	96.87	49.22	74.08	30.33	6.36	23.10	6.22	37.53	11.71	98.94	99.92	404.08	434.42	108.47	124.82
Spleen	3.2%	98.69	97.99	82.38	69.89	78.75	72.79	67.39	55.42	80.52	71.31	99.40	98.89	286.23	309.00	54.85	74.29
Stomach	5.0%	94.89	95.28	51.48	57.70	63.62	41.38	39.77	31.75	56.91	48.20	96.64	98.30	513.78	410.70	122.87	137.24
Colon	11.8%	92.95	91.60	72.26	68.61	72.68	63.01	56.83	48.91	72.47	65.69	95.92	95.78	446.84	433.33	77.08	90.38
Small Intestine	15.5%	94.47	93.27	77.92	74.35	91.14	88.21	72.43	67.63	84.01	80.69	95.10	94.23	440.6	459.11	56.50	66.05
Liver	19.7%	93.42	91.47	82.24	77.88	70.46	58.58	61.16	50.23	75.99	66.87	97.38	97.13	368.69	475.25	71.06	131.55
Abdominal Wall	26.2%	92.48	91.39	91.41	89.04	81.10	79.44	75.35	72.36	85.95	83.97	96.98	96.13	329.73	358.67	38.37	43.32

Supplementary Table 2: Performance of binary organ segmentation models trained with class weights adapted to organ size and tested on organ-specific test data. Models were trained on organ-specific (O) training data including only class-positive images or supplemented (S) training data including class-positive and class-negative examples. All models were tested on organ-specific test data. Abbreviations: Average Symmetric Surface Distance (ASSD), Hausdorff Distance (HD), Intersection-over-Union (IoU), Organ-specific training dataset (O), Supplemented training dataset (S).

Organ	Organ Size (% Foreground)	Accuracy		Precision		Recall		IoU		F1 Score		Specificity		HD		ASSD	
		O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S
Ureter	1.2%	98.81	98.53	45.51	32.22	41.52	30.65	27.73	18.63	43.43	31.41	99.45	99.28	405.10	436.90	169.96	180.67
Intestinal Veins	1.3%	99.35	99.07	70.02	58.03	79.29	79.52	59.19	50.48	74.36	67.10	99.59	99.31	271.91	251.89	46.14	46.83
Pancreas	2.7%	98.13	98.13	65.40	76.91	61.38	41.27	46.33	36.72	63.33	53.71	99.12	99.67	343.57	340.13	90.85	91.49
Inferior Mesenteric Artery	2.8%	97.95	97.54	63.05	51.39	45.77	44.42	36.09	31.28	53.04	47.65	99.31	98.91	263.44	301.03	67.83	70.27
Vesicular Glands	2.9%	95.83	96.07	36.58	37.66	37.98	31.13	22.90	20.54	37.27	34.09	97.78	98.26	399.46	406.33	98.11	107.02
Spleen	3.2%	98.11	97.83	68.05	64.89	84.77	79.93	60.64	55.80	75.50	71.63	98.59	98.46	296.75	361.33	57.94	74.66
Stomach	5.0%	94.09	91.60	46.24	36.65	70.65	80.04	38.79	33.58	55.90	50.28	95.40	92.25	540.69	656.05	133.18	158.10
Colon	11.8%	92.69	89.71	68.79	57.16	78.25	77.39	57.75	48.98	73.22	65.76	94.80	91.51	498.50	500.49	82.19	86.24
Small Intestine	15.5%	95.33	91.25	83.50	65.90	88.14	93.48	75.06	63.00	85.76	77.30	96.70	90.83	355.14	538.04	47.13	80.89
Liver	19.7%	92.95	91.24	75.07	75.10	77.91	60.46	61.89	50.36	76.46	66.99	95.54	96.55	424.64	447.13	73.86	109.35
Abdominal Wall	26.2%	92.60	91.26	88.38	88.93	85.09	79.03	76.53	71.95	86.70	83.69	95.57	96.11	350.55	363.21	38.39	44.88