

Supplementary Material

We report additional quantitative results across three datasets under different occlusion generation strategies and prompt settings. Tables 1–4 present results on CVC-300, Tables 5–8 on CVC-ColonDB, and Tables 9–12 on ETIS-LaribPolypDB. Each table reports Dice Similarity Coefficient (DSC, \uparrow) and 95th percentile Hausdorff Distance (HD95, \downarrow) under four severity levels (Clean, Low, Medium, High), evaluated across three modes: *Full (Amodal)*, *Visible Only*, and *Invisible (Occluded)*. Results consistently show that performance degrades with increasing occlusion severity, with the Invisible mode being the most challenging as models must infer occluded regions without direct visual evidence. Box prompts generally yield stronger results than Point prompts across all settings.

Table 1. Segmentation performance on CVC-300 under Surgical Tool occlusion using Box prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow
Full (Amodal)	SAM	0.93	10.05	0.85	38.30	0.75	62.17	0.58	73.95
	SAM 2	0.91	8.48	<u>0.87</u>	<u>35.34</u>	0.77	61.64	<u>0.69</u>	65.53
	SAM 3	0.95	4.98	0.87	35.51	0.78	62.42	0.62	<u>60.11</u>
	MedSAM	0.74	24.43	0.63	53.34	0.60	68.74	0.64	64.27
	SAM-Med2D	0.90	12.78	0.84	36.47	0.70	66.66	0.55	75.47
	MedSAM2	0.93	7.40	0.91	22.85	0.83	42.28	0.78	51.21
	MedSAM3	<u>0.94</u>	<u>6.69</u>	0.85	38.06	<u>0.79</u>	<u>61.30</u>	0.68	62.09
Visible Only	SAM	0.93	10.05	0.89	22.67	<u>0.91</u>	17.72	0.39	61.41
	SAM 2	0.91	8.48	<u>0.91</u>	<u>16.27</u>	0.89	10.92	0.66	<u>31.65</u>
	SAM 3	0.95	4.98	<u>0.91</u>	16.40	0.93	<u>12.80</u>	0.72	27.53
	MedSAM	0.74	24.43	0.67	48.50	0.58	54.28	0.30	80.01
	SAM-Med2D	0.90	12.78	0.89	23.57	0.78	34.57	0.35	71.11
	MedSAM2	0.93	7.40	0.94	15.41	0.87	22.11	<u>0.69</u>	33.96
	MedSAM3	<u>0.94</u>	<u>6.69</u>	0.88	23.72	0.89	24.84	0.66	40.74
Invisible (Occluded)	SAM	–	–	0.04	62.98	0.00	105.32	0.59	47.30
	SAM 2	–	–	0.12	43.98	0.19	61.38	0.57	50.49
	SAM 3	–	–	0.05	68.19	0.06	71.62	0.34	65.63
	MedSAM	–	–	<u>0.39</u>	<u>32.57</u>	<u>0.59</u>	34.51	0.82	29.25
	SAM-Med2D	–	–	0.07	53.41	0.30	67.90	<u>0.66</u>	48.26
	MedSAM2	–	–	0.49	25.90	0.66	<u>34.80</u>	0.82	<u>37.88</u>
	MedSAM3	–	–	0.12	33.08	0.21	48.52	0.45	54.07

Table 2. Segmentation performance on CVC-300 under Cutout occlusion using Box prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓
Full (Amodal)	SAM	0.93	10.05	0.89	17.48	0.69	35.73	0.61	36.59
	SAM 2	0.91	8.48	0.92	15.81	0.77	31.93	0.68	31.15
	SAM 3	0.95	4.98	0.94	6.95	0.89	<u>19.67</u>	0.87	27.55
	MedSAM	0.74	24.43	0.68	30.31	0.60	37.29	0.64	32.04
	SAM-Med2D	0.90	12.78	0.87	20.81	0.75	30.07	0.68	32.14
	MedSAM2	0.93	7.40	0.91	14.18	0.82	24.12	0.77	<u>25.65</u>
	MedSAM3	<u>0.94</u>	<u>6.69</u>	<u>0.93</u>	<u>9.47</u>	<u>0.86</u>	16.13	<u>0.84</u>	23.00
Visible Only	SAM	0.93	10.05	0.92	13.11	0.74	19.96	0.78	20.05
	SAM 2	0.91	8.48	0.92	8.93	<u>0.84</u>	16.34	0.92	6.70
	SAM 3	0.95	4.98	<u>0.91</u>	14.09	0.81	23.55	<u>0.79</u>	25.63
	MedSAM	0.74	24.43	0.63	25.56	0.35	33.63	0.25	34.01
	SAM-Med2D	0.90	12.78	0.86	15.83	0.75	19.91	0.75	20.51
	MedSAM2	0.93	7.40	<u>0.91</u>	<u>10.57</u>	0.85	<u>17.13</u>	<u>0.79</u>	<u>19.85</u>
	MedSAM3	<u>0.94</u>	<u>6.69</u>	0.90	18.41	0.75	31.61	0.75	28.30
Invisible (Occluded)	SAM	-	-	0.15	17.77	0.24	27.74	0.24	20.21
	SAM 2	-	-	0.54	10.44	0.24	25.36	0.21	44.04
	SAM 3	-	-	0.77	6.92	0.78	<u>10.29</u>	<u>0.78</u>	<u>13.45</u>
	MedSAM	-	-	<u>0.82</u>	3.58	0.90	8.60	0.86	10.46
	SAM-Med2D	-	-	0.39	10.51	0.46	16.71	0.45	26.40
	MedSAM2	-	-	0.50	7.74	0.49	16.66	0.51	30.44
	MedSAM3	-	-	0.88	<u>6.00</u>	<u>0.80</u>	11.89	0.72	19.87

Table 3. Segmentation performance on CVC-300 under Surgical Tool occlusion using Point prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓
Full (Amodal)	SAM	0.49	165.54	0.23	468.78	0.18	517.32	0.39	206.37
	SAM 2	<u>0.83</u>	44.03	0.66	130.22	0.59	144.94	0.49	130.24
	SAM 3	0.93	6.71	0.73	105.00	0.52	112.44	0.50	129.06
	MedSAM	0.02	89.81	0.00	180.03	0.01	177.41	0.02	150.37
	SAM-Med2D	0.68	40.67	0.57	79.31	0.38	<u>99.08</u>	0.51	99.78
	MedSAM2	0.77	25.96	0.66	83.86	<u>0.60</u>	86.40	0.51	<u>111.29</u>
	MedSAM3	0.93	<u>8.24</u>	<u>0.73</u>	120.35	0.68	103.49	<u>0.49</u>	120.83
Visible Only	SAM	0.49	165.54	0.37	429.25	0.24	533.19	0.06	607.70
	SAM 2	0.83	44.03	<u>0.81</u>	100.95	0.78	104.27	0.53	210.94
	SAM 3	0.93	6.71	0.82	90.20	<u>0.71</u>	109.89	0.62	155.72
	MedSAM	0.02	89.81	0.00	183.98	0.04	162.40	0.01	148.48
	SAM-Med2D	0.68	40.67	0.63	73.62	0.50	<u>101.63</u>	0.48	<u>122.43</u>
	MedSAM2	0.77	25.96	0.77	<u>75.87</u>	0.63	68.00	0.49	97.05
	MedSAM3	0.93	<u>8.24</u>	0.79	122.56	<u>0.71</u>	132.67	<u>0.58</u>	158.50
Invisible (Occluded)	SAM	-	-	0.00	498.83	0.00	587.13	0.01	616.51
	SAM 2	-	-	0.00	220.57	0.00	180.08	0.00	251.46
	SAM 3	-	-	<u>0.09</u>	80.90	<u>0.35</u>	<u>85.45</u>	0.52	90.40
	MedSAM	-	-	<u>0.02</u>	141.66	0.04	153.79	0.04	151.61
	SAM-Med2D	-	-	0.05	<u>81.99</u>	0.22	100.70	0.67	83.45
	MedSAM2	-	-	0.05	181.30	0.22	108.57	0.38	110.86
	MedSAM3	-	-	0.18	83.47	0.42	83.15	<u>0.66</u>	<u>86.52</u>

Table 4. Segmentation performance on CVC-300 under Cutout occlusion using Point prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓
Full (Amodal)	SAM	0.49	165.54	0.29	227.06	0.34	146.29	0.38	122.88
	SAM 2	<u>0.83</u>	44.03	0.78	33.04	<u>0.54</u>	55.09	0.52	50.47
	SAM 3	0.93	6.71	0.80	<u>30.49</u>	0.75	41.28	0.68	<u>46.37</u>
	MedSAM	0.02	89.81	0.03	81.26	0.01	82.60	0.02	72.76
	SAM-Med2D	0.68	40.67	0.54	38.21	0.45	<u>45.54</u>	0.46	51.26
	MedSAM2	0.77	25.96	0.83	20.28	0.52	71.16	0.58	45.31
	MedSAM3	0.93	<u>8.24</u>	<u>0.82</u>	34.66	0.75	81.03	<u>0.66</u>	73.78
Visible Only	SAM	0.49	165.54	0.29	248.05	0.28	206.11	0.06	324.66
	SAM 2	<u>0.83</u>	44.03	<u>0.85</u>	<u>32.07</u>	0.64	69.77	<u>0.69</u>	68.06
	SAM 3	0.93	6.71	0.82	<u>37.83</u>	<u>0.66</u>	69.64	0.78	33.73
	MedSAM	0.02	89.81	0.02	93.28	0.01	88.15	0.01	72.55
	SAM-Med2D	0.68	40.67	0.58	35.65	0.45	55.44	0.48	50.48
	MedSAM2	0.77	25.96	0.87	12.12	0.50	<u>56.18</u>	0.59	<u>41.83</u>
	MedSAM3	0.93	<u>8.24</u>	0.80	52.75	0.68	82.84	0.62	78.52
Invisible (Occluded)	SAM	-	-	0.55	<u>5.85</u>	<u>0.60</u>	23.91	0.54	50.08
	SAM 2	-	-	0.18	21.06	0.51	26.73	0.48	49.48
	SAM 3	-	-	<u>0.64</u>	5.38	<u>0.60</u>	<u>20.38</u>	0.50	50.49
	MedSAM	-	-	0.20	11.47	0.03	48.81	0.02	61.91
	SAM-Med2D	-	-	0.34	5.91	0.57	24.59	0.49	47.09
	MedSAM2	-	-	0.19	5.94	0.32	30.42	0.50	31.28
	MedSAM3	-	-	0.81	7.37	0.74	18.44	<u>0.53</u>	<u>41.24</u>

Table 5. Segmentation performance on CVC-ColonDB under Surgical Tool occlusion using Box prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓
Full (Amodal)	SAM	0.88	20.48	0.80	86.41	0.60	110.62	0.57	95.72
	SAM 2	<u>0.91</u>	16.24	0.84	67.70	0.73	80.37	0.65	71.77
	SAM 3	0.93	13.51	0.87	<u>62.71</u>	0.76	<u>72.79</u>	0.63	71.57
	MedSAM	0.71	36.73	0.66	88.09	0.57	90.08	0.61	77.75
	SAM-Med2D	0.85	28.24	0.76	77.80	0.57	97.07	0.54	79.34
	MedSAM2	<u>0.91</u>	14.73	0.87	54.72	0.82	62.38	0.76	59.00
	MedSAM3	<u>0.91</u>	<u>14.43</u>	<u>0.85</u>	63.42	<u>0.77</u>	73.72	<u>0.66</u>	<u>66.89</u>
Visible Only	SAM	0.88	20.48	0.84	62.35	0.63	73.93	0.50	67.63
	SAM 2	<u>0.91</u>	16.24	<u>0.89</u>	44.30	0.82	<u>41.57</u>	<u>0.73</u>	<u>43.37</u>
	SAM 3	0.93	13.51	0.91	35.73	0.88	36.50	0.75	40.14
	MedSAM	0.71	36.73	0.66	79.30	0.46	77.80	0.28	75.33
	SAM-Med2D	0.85	28.24	0.80	60.37	0.54	71.63	0.39	66.02
	MedSAM2	<u>0.91</u>	14.73	0.87	46.24	0.75	65.33	0.59	64.57
	MedSAM3	<u>0.91</u>	<u>14.43</u>	0.88	<u>44.00</u>	<u>0.84</u>	43.42	0.72	46.67
Invisible (Occluded)	SAM	-	-	0.05	101.92	0.29	113.93	0.50	78.48
	SAM 2	-	-	0.11	70.81	0.28	89.12	0.41	80.17
	SAM 3	-	-	0.08	65.79	0.17	82.78	0.30	73.83
	MedSAM	-	-	0.49	<u>40.17</u>	<u>0.64</u>	<u>50.65</u>	0.79	37.08
	SAM-Med2D	-	-	0.13	73.99	0.42	78.97	0.63	56.76
	MedSAM2	-	-	<u>0.48</u>	36.26	0.72	48.13	<u>0.74</u>	<u>52.83</u>
	MedSAM3	-	-	0.22	45.62	0.30	70.61	0.41	63.79

Table 6. Segmentation performance on CVC-ColonDB under Cutout occlusion using Box prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow
Full (Amodal)	SAM	0.88	20.48	0.82	30.02	0.66	49.17	0.58	59.09
	SAM 2	<u>0.91</u>	16.24	0.88	23.22	0.77	39.30	0.66	45.02
	SAM 3	0.93	13.51	0.91	14.53	0.90	<u>29.74</u>	0.68	34.97
	MedSAM	0.71	36.73	0.65	40.02	0.58	45.92	0.65	42.88
	SAM-Med2D	0.85	28.24	0.81	31.89	0.71	41.60	0.61	46.97
	MedSAM2	<u>0.91</u>	14.73	<u>0.89</u>	22.26	0.80	37.80	<u>0.72</u>	40.93
	MedSAM3	<u>0.91</u>	<u>14.43</u>	<u>0.89</u>	<u>19.39</u>	<u>0.86</u>	28.95	0.79	<u>37.34</u>
Visible Only	SAM	0.88	20.48	0.83	21.98	0.69	29.52	0.64	30.41
	SAM 2	<u>0.91</u>	16.24	0.90	<u>16.72</u>	0.86	16.43	<u>0.84</u>	17.13
	SAM 3	0.93	13.51	<u>0.89</u>	9.59	0.76	55.54	0.90	45.00
	MedSAM	0.71	36.73	0.59	37.32	0.40	42.58	0.31	45.25
	SAM-Med2D	0.85	28.24	0.80	27.77	0.70	31.48	0.63	32.65
	MedSAM2	<u>0.91</u>	14.73	<u>0.89</u>	17.33	<u>0.84</u>	<u>21.15</u>	0.79	<u>26.79</u>
	MedSAM3	<u>0.91</u>	<u>14.43</u>	0.86	23.04	0.81	27.38	0.74	33.74
Invisible (Occluded)	SAM	-	-	0.19	17.35	0.25	32.21	0.34	38.71
	SAM 2	-	-	0.23	24.89	0.18	55.05	0.21	70.73
	SAM 3	-	-	<u>0.68</u>	6.68	0.81	31.37	0.21	<u>31.11</u>
	MedSAM	-	-	0.66	<u>12.94</u>	0.81	18.71	0.84	23.03
	SAM-Med2D	-	-	0.37	18.60	0.43	39.72	0.41	48.17
	MedSAM2	-	-	0.35	19.00	0.39	33.40	0.45	42.21
	MedSAM3	-	-	0.71	13.52	<u>0.68</u>	<u>22.44</u>	<u>0.63</u>	33.60

Table 7. Segmentation performance on CVC-ColonDB under Surgical Tool occlusion using Point prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow
Full (Amodal)	SAM	0.32	248.89	0.26	496.00	0.24	439.15	0.34	305.62
	SAM 2	0.74	72.05	0.56	241.01	0.43	248.51	0.45	177.26
	SAM 3	<u>0.78</u>	<u>50.32</u>	<u>0.65</u>	<u>132.18</u>	0.49	170.44	0.43	155.95
	MedSAM	0.02	133.46	0.02	265.20	0.02	252.97	0.02	194.77
	SAM-Med2D	0.53	67.91	0.40	155.28	0.35	172.58	0.45	133.59
	MedSAM2	0.73	51.94	0.64	114.44	<u>0.54</u>	142.81	<u>0.50</u>	<u>131.87</u>
	MedSAM3	0.81	47.77	0.73	134.29	0.56	<u>170.06</u>	0.53	131.18
Visible Only	SAM	0.32	248.89	0.31	476.20	0.26	507.95	0.14	583.24
	SAM 2	0.74	72.05	0.65	202.48	0.58	261.97	0.53	283.68
	SAM 3	<u>0.78</u>	<u>50.32</u>	<u>0.68</u>	140.88	<u>0.64</u>	171.77	<u>0.59</u>	189.79
	MedSAM	0.02	133.46	0.02	254.89	0.02	247.95	0.01	198.21
	SAM-Med2D	0.53	67.91	0.44	154.55	0.45	167.25	0.48	<u>157.12</u>
	MedSAM2	0.73	51.94	0.67	109.97	0.59	130.40	0.52	143.27
	MedSAM3	0.81	47.77	0.74	<u>136.15</u>	0.70	<u>142.54</u>	0.62	158.74
Invisible (Occluded)	SAM	-	-	0.00	630.90	0.00	586.58	0.00	620.33
	SAM 2	-	-	0.00	380.16	0.01	345.82	0.00	329.40
	SAM 3	-	-	0.09	83.69	<u>0.33</u>	107.34	0.40	107.33
	MedSAM	-	-	0.01	205.22	<u>0.02</u>	232.07	0.02	197.08
	SAM-Med2D	-	-	0.04	60.50	0.23	<u>101.50</u>	<u>0.55</u>	84.84
	MedSAM2	-	-	0.19	<u>72.73</u>	0.41	99.85	0.57	109.25
	MedSAM3	-	-	<u>0.16</u>	79.03	0.31	101.84	0.53	<u>104.77</u>

Table 8. Segmentation performance on CVC-ColonDB under Cutout occlusion using Point prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓
Full (Amodal)	SAM	0.32	248.89	0.27	238.48	0.29	189.87	0.46	106.36
	SAM 2	0.74	72.05	<u>0.66</u>	73.48	0.55	80.03	0.51	<u>89.63</u>
	SAM 3	<u>0.78</u>	<u>50.32</u>	0.58	24.94	0.50	90.58	<u>0.58</u>	154.00
	MedSAM	0.02	133.46	0.02	125.80	0.02	116.01	0.02	118.44
	SAM-Med2D	0.53	67.91	0.43	71.74	0.35	79.62	0.39	81.40
	MedSAM2	0.73	51.94	0.65	60.75	<u>0.58</u>	65.98	0.49	105.14
	MedSAM3	0.81	47.77	0.78	<u>55.53</u>	0.72	<u>73.25</u>	0.63	96.59
	Visible Only	SAM	0.32	248.89	0.26	258.86	0.23	278.66	0.21
SAM 2		0.74	72.05	0.70	75.53	<u>0.69</u>	85.51	<u>0.66</u>	93.40
SAM 3		<u>0.78</u>	<u>50.32</u>	0.81	7.47	0.43	282.65	0.89	158.00
MedSAM		0.02	133.46	0.02	128.57	0.02	131.78	0.01	122.71
SAM-Med2D		0.53	67.91	0.46	69.85	0.40	87.16	0.43	90.27
MedSAM2		0.73	51.94	0.67	<u>61.95</u>	0.68	56.83	0.61	<u>73.46</u>
MedSAM3		0.81	47.77	<u>0.74</u>	62.34	0.70	<u>66.45</u>	<u>0.66</u>	73.24
Invisible (Occluded)		SAM	-	-	0.43	20.41	0.41	42.18	0.55
	SAM 2	-	-	0.11	35.37	0.27	50.86	<u>0.52</u>	59.21
	SAM 3	-	-	<u>0.55</u>	12.18	<u>0.42</u>	14.00	0.02	138.00
	MedSAM	-	-	0.04	28.53	0.04	59.64	0.02	88.42
	SAM-Med2D	-	-	0.29	10.88	0.38	<u>32.80</u>	0.46	<u>54.92</u>
	MedSAM2	-	-	0.14	29.67	0.17	45.13	0.40	56.12
	MedSAM3	-	-	0.57	17.42	0.52	35.73	0.55	52.44

Table 9. Segmentation performance on ETIS-LaribPolypDB under Surgical Tool occlusion using Box prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓	DSC ↑	HD95 ↓
Full (Amodal)	SAM	0.91	33.13	0.81	45.30	0.66	84.53	0.55	91.15
	SAM 2	0.91	33.18	<u>0.84</u>	48.94	0.74	70.69	0.62	82.04
	SAM 3	0.94	22.15	<u>0.84</u>	32.71	<u>0.76</u>	<u>66.51</u>	0.60	<u>66.48</u>
	MedSAM	0.78	47.40	0.72	46.89	0.62	71.77	0.59	76.70
	SAM-Med2D	0.84	52.78	0.78	46.87	0.61	77.94	0.50	87.52
	MedSAM2	0.91	30.57	0.87	<u>34.29</u>	0.81	51.99	0.77	56.57
	MedSAM3	<u>0.92</u>	<u>29.13</u>	0.87	34.55	<u>0.76</u>	69.65	<u>0.68</u>	70.43
	Visible Only	SAM	0.91	33.13	0.84	38.60	0.78	47.37	0.55
SAM 2		0.91	33.18	0.88	36.96	<u>0.86</u>	<u>33.84</u>	0.72	<u>49.82</u>
SAM 3		0.94	22.15	<u>0.89</u>	18.82	0.89	26.92	0.80	29.95
MedSAM		0.78	47.40	0.76	39.70	0.60	53.23	0.39	66.79
SAM-Med2D		0.84	52.78	0.82	36.60	0.65	52.55	0.49	63.00
MedSAM2		0.91	30.57	<u>0.89</u>	29.51	0.80	48.56	0.64	64.14
MedSAM3		<u>0.92</u>	<u>29.13</u>	0.91	<u>22.60</u>	0.84	42.67	<u>0.74</u>	50.56
Invisible (Occluded)		SAM	-	-	0.03	75.04	0.16	115.42	0.44
	SAM 2	-	-	0.12	47.62	0.16	73.67	0.39	75.98
	SAM 3	-	-	0.04	51.98	0.11	75.52	0.18	76.34
	MedSAM	-	-	<u>0.22</u>	43.00	<u>0.55</u>	41.49	<u>0.69</u>	48.01
	SAM-Med2D	-	-	0.10	54.94	0.30	65.36	0.48	67.48
	MedSAM2	-	-	0.36	29.42	0.56	<u>43.86</u>	0.71	<u>51.71</u>
	MedSAM3	-	-	0.15	<u>33.02</u>	0.27	63.41	0.44	63.85

Table 10. Segmentation performance on ETIS-LaribPolypDB under Cutout occlusion using Box prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow
Full (Amodal)	SAM	0.91	33.13	0.82	50.09	0.68	79.46	0.63	92.08
	SAM 2	0.91	33.18	0.88	43.65	0.75	70.49	0.66	79.74
	SAM 3	0.94	22.15	0.92	31.75	0.86	<u>60.73</u>	0.77	69.34
	MedSAM	0.78	47.40	0.70	55.99	0.67	69.33	0.67	<u>68.51</u>
	SAM-Med2D	0.84	52.78	0.79	57.14	0.72	71.58	0.62	81.02
	MedSAM2	0.91	30.57	0.87	42.51	<u>0.80</u>	66.70	0.67	80.92
	MedSAM3	<u>0.92</u>	<u>29.13</u>	<u>0.91</u>	<u>31.91</u>	0.86	57.10	<u>0.76</u>	64.38
Visible Only	SAM	0.91	33.13	0.82	39.68	0.67	47.75	0.70	44.71
	SAM 2	0.91	33.18	0.90	<u>33.61</u>	<u>0.84</u>	<u>37.76</u>	0.81	38.29
	SAM 3	0.94	22.15	0.90	28.91	0.83	43.71	0.81	<u>41.42</u>
	MedSAM	0.78	47.40	0.66	53.10	0.50	67.51	0.37	76.33
	SAM-Med2D	0.84	52.78	0.78	50.28	0.68	59.32	0.64	56.64
	MedSAM2	0.91	30.57	<u>0.88</u>	34.04	0.87	31.47	<u>0.79</u>	49.93
	MedSAM3	<u>0.92</u>	<u>29.13</u>	<u>0.88</u>	37.20	0.80	52.04	0.77	46.84
Invisible (Occluded)	SAM	-	-	0.16	38.30	0.29	61.02	0.33	84.98
	SAM 2	-	-	0.21	42.73	0.22	90.07	0.27	108.33
	SAM 3	-	-	0.57	19.87	0.63	46.34	0.52	<u>52.22</u>
	MedSAM	-	-	<u>0.59</u>	24.60	0.82	37.02	0.84	37.10
	SAM-Med2D	-	-	0.41	28.75	0.51	61.39	0.39	75.26
	MedSAM2	-	-	0.27	30.30	0.28	63.97	0.31	82.63
	MedSAM3	-	-	0.67	<u>21.27</u>	<u>0.67</u>	<u>44.66</u>	<u>0.57</u>	52.43

Table 11. Segmentation performance on ETIS-LaribPolypDB under Surgical Tool occlusion using Point prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow
Full (Amodal)	SAM	0.46	405.77	0.28	503.74	0.28	414.99	0.32	296.23
	SAM 2	0.69	225.01	0.57	220.11	0.44	245.45	0.40	201.71
	SAM 3	<u>0.75</u>	125.54	<u>0.66</u>	133.30	<u>0.48</u>	161.10	<u>0.44</u>	164.78
	MedSAM	0.02	232.68	0.03	180.45	0.03	175.92	0.02	192.56
	SAM-Med2D	0.67	112.44	0.58	102.98	0.42	<u>132.24</u>	<u>0.44</u>	134.33
	MedSAM2	0.71	<u>123.13</u>	0.61	<u>125.50</u>	0.56	115.28	0.47	<u>140.05</u>
	MedSAM3	0.79	142.88	0.70	155.37	0.56	152.14	0.47	146.07
Visible Only	SAM	0.46	405.77	0.35	457.31	0.37	442.35	0.27	498.70
	SAM 2	0.69	225.01	0.66	203.64	0.65	214.55	0.56	287.09
	SAM 3	<u>0.75</u>	125.54	0.76	<u>113.28</u>	<u>0.66</u>	<u>119.13</u>	<u>0.57</u>	197.49
	MedSAM	0.02	232.68	0.02	181.17	0.02	203.16	0.03	179.35
	SAM-Med2D	0.67	112.44	0.65	81.94	0.55	132.59	0.55	<u>165.54</u>
	MedSAM2	0.71	<u>123.13</u>	0.63	117.25	0.64	96.70	0.51	142.46
	MedSAM3	0.79	142.88	<u>0.73</u>	153.81	0.69	154.72	0.58	216.75
Invisible (Occluded)	SAM	-	-	0.00	546.40	0.00	514.11	0.00	541.68
	SAM 2	-	-	0.00	292.00	0.00	301.52	0.00	320.93
	SAM 3	-	-	0.05	<u>83.38</u>	0.21	110.88	0.35	104.73
	MedSAM	-	-	0.00	164.57	0.02	189.95	0.03	170.84
	SAM-Med2D	-	-	<u>0.06</u>	83.73	<u>0.22</u>	94.73	0.52	<u>101.17</u>
	MedSAM2	-	-	0.02	210.20	0.11	158.52	0.15	167.37
	MedSAM3	-	-	0.13	80.06	0.24	<u>95.62</u>	<u>0.47</u>	101.12

Table 12. Segmentation performance on ETIS-LaribPolypDB under Cutout occlusion using Point prompts across clean, low, medium, and high severity levels. **Bold** and underline denote the best and second best per column, respectively.

Eval Mode	Model	Clean		Low		Medium		High	
		DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow	DSC \uparrow	HD95 \downarrow
Full (Amodal)	SAM	0.46	405.77	0.40	382.44	0.32	377.71	0.45	209.90
	SAM 2	0.69	225.01	0.65	165.76	0.48	193.81	0.49	188.75
	SAM 3	<u>0.75</u>	125.54	<u>0.78</u>	98.36	0.64	140.81	0.61	<u>153.30</u>
	MedSAM	0.02	232.68	0.03	198.13	0.03	207.02	0.02	226.62
	SAM-Med2D	0.67	112.44	0.58	<u>103.03</u>	0.46	<u>139.07</u>	0.49	136.29
	MedSAM2	0.71	<u>123.13</u>	0.66	122.91	<u>0.54</u>	121.36	0.41	222.15
	MedSAM3	0.79	142.88	0.79	133.93	0.64	215.01	<u>0.53</u>	247.12
Visible Only	SAM	0.46	405.77	0.34	463.27	0.33	532.38	0.37	454.53
	SAM 2	0.69	225.01	0.67	201.99	0.63	235.22	<u>0.63</u>	240.06
	SAM 3	<u>0.75</u>	125.54	0.79	102.31	0.74	<u>130.19</u>	0.71	120.16
	MedSAM	0.02	232.68	0.02	210.44	0.02	202.64	0.01	204.14
	SAM-Med2D	0.67	112.44	0.60	<u>104.23</u>	0.52	138.13	0.52	<u>139.44</u>
	MedSAM2	0.71	<u>123.13</u>	0.70	114.31	<u>0.68</u>	91.83	0.59	179.20
	MedSAM3	0.79	142.88	<u>0.76</u>	124.01	0.63	228.50	0.60	194.87
Invisible (Occluded)	SAM	-	-	0.28	48.86	0.30	80.21	0.44	97.80
	SAM 2	-	-	0.13	51.98	0.27	93.70	0.41	120.41
	SAM 3	-	-	<u>0.42</u>	28.96	0.39	65.40	<u>0.45</u>	95.65
	MedSAM	-	-	0.04	56.35	0.07	103.81	0.02	126.89
	SAM-Med2D	-	-	0.39	17.24	<u>0.50</u>	<u>62.61</u>	0.47	86.63
	MedSAM2	-	-	0.08	44.50	0.12	84.81	0.23	105.53
	MedSAM3	-	-	0.59	<u>28.54</u>	0.54	52.76	<u>0.45</u>	<u>90.78</u>