

Method	Type	K	PH2	ISIC	Optical	Vessel	Xray
Random	Data	2	79.6	61	19.5	26.1	73.6
		4	79.8	63.1	46.1	41	77.8
		8	80.1	67.9	64.4	59	82.1
		16	80.3	69.1	73	62.3	86.4
		32	81.5	70.5	/	/	87.2
TopK	Data	2	83.9	62.5	23.9	32	76.1
		4	84	67.2	55	44	79
		8	84.2	72.3	65.5	61	84.3
		16	85.1	74.5	73.5	64.5	87.7
		32	85.8	76.5	/	/	88.3
MVPS (Gain Over TopK)	Data	2	<b>86.2</b> ↑ 2.3	<b>66.8</b> ↑ 4.3	<b>31.4</b> ↑ 7.5	<b>32.0</b> ↑ 0.0	<b>78.6</b> ↑ 2.5
		4	<b>88.2</b> ↑ 4.2	<b>72.8</b> ↑ 5.1	<b>59.9</b> ↑ 4.9	<b>48.6</b> ↑ 4.6	<b>82.6</b> ↑ 3.6
		8	<b>88.5</b> ↑ 4.3	<b>76.2</b> ↑ 3.6	<b>73.6</b> ↑ 8.1	<b>71.0</b> ↑ 10.0	<b>86.2</b> ↑ 1.9
		16	<b>89.1</b> ↑ 4.0	<b>76.5</b> ↑ 2.0	<b>79.9</b> ↑ 6.4	<b>72.6</b> ↑ 7.1	<b>88.1</b> ↑ 0.4
		32	<b>89.5</b> ↑ 3.7	<b>77.9</b> ↑ 1.4	/	/	<b>89.1</b> ↑ 0.8
+TTA (Gain Over TopK)	Data	2	86.2	66.8	31.4	32	79.1
		4	88.2	72.8	59.9	48.6	85.4
		8	88.5	76.2	73.6	71	88.2
		16	90.9	77.4	81.1	73.2	88.5
		32	90.9	78.3	/	/	89.5
Lora	Model	2	83.8	75.2	65.8	45.3	79.3
		4	85.2	75.6	73.4	61.2	83.8
		8	86.3	78.2	79.3	72.5	88.8
		16	89.8	80.6	86.9	74.6	91.1
		32	91.5	82.3	/	/	93.3
+MVPS (Gain Over Lora)	Combined	2	<b>89.5</b> ↑ 5.7	<b>76.1</b> ↑ 0.9	<b>66.3</b> ↑ 0.5	<b>51.8</b> ↑ 6.5	<b>82.4</b> ↑ 3.1
		4	<b>89.9</b> ↑ 4.7	<b>78.5</b> ↑ 2.9	<b>76.5</b> ↑ 3.1	<b>68.6</b> ↑ 7.4	<b>85.9</b> ↑ 2.1
		8	<b>90.3</b> ↑ 4.0	<b>79.4</b> ↑ 1.2	<b>81.7</b> ↑ 2.4	<b>75.3</b> ↑ 2.8	<b>89.8</b> ↑ 1.0
		16	<b>91.2</b> ↑ 1.4	<b>83.2</b> ↑ 2.6	<b>89.7</b> ↑ 2.8	<b>75.7</b> ↑ 1.1	<b>92.1</b> ↑ 1.0
		32	<b>93.3</b> ↑ 0.8	<b>92.3</b> ↑ 2.8	/	/	<b>95.1</b> ↑ 1.8
Supervised -	-	96.4	94.8	98.1	86.8	96.6	

Table 1: DICE Score Over Different Datasets and K number of prompts.

Method	Type	K	PH2	ISIC	Xray	K	Optical	Vessel
MVPS	Data	4	<b>88.2</b> $\uparrow$ 4.2	<b>72.3</b> $\uparrow$ 5.1	<b>82.6</b> $\uparrow$ 3.6	4	<b>59.9</b> $\uparrow$ 4.9	<b>48.6</b> $\uparrow$ 4.6
		8	<b>88.5</b> $\uparrow$ 4.3	<b>76.2</b> $\uparrow$ 3.6	<b>86.2</b> $\uparrow$ 1.9	8	<b>73.6</b> $\uparrow$ 8.1	<b>71.0</b> $\uparrow$ 10.0
		32	<b>89.5</b> $\uparrow$ 3.7	<b>77.9</b> $\uparrow$ 1.4	<b>89.1</b> $\uparrow$ 0.8	16	<b>79.9</b> $\uparrow$ 6.4	<b>72.6</b> $\uparrow$ 7.1
MVPS+TTA	Data	4	<b>88.2</b> $\uparrow$ 4.2	<b>72.8</b> $\uparrow$ 5.6	<b>85.4</b> $\uparrow$ 6.4	4	<b>59.9</b> $\uparrow$ 4.9	<b>48.6</b> $\uparrow$ 4.6
		8	<b>88.5</b> $\uparrow$ 4.3	<b>76.2</b> $\uparrow$ 4.0	<b>88.2</b> $\uparrow$ 3.9	8	<b>73.6</b> $\uparrow$ 8.1	<b>71.0</b> $\uparrow$ 10.0
		32	<b>90.9</b> $\uparrow$ 5.1	<b>78.3</b> $\uparrow$ 1.8	<b>89.5</b> $\uparrow$ 1.2	16	<b>81.1</b> $\uparrow$ 7.6	<b>73.2</b> $\uparrow$ 7.7

Table 2: Ablation: Test Time Adaptation

	PH2	ISIC	Optical	Disc	Vessel	X-Ray
<b>K=2</b>	85.94	66.56	30.86	31.57	78.82	
<b>K=4</b>	88.17	72.69	59.77	47.97	81.95	
<b>K=8</b>	88.31	76.18	73.43	70.65	85.88	
<b>K=16</b>	90.66	76.96	79.22	/	88.03	
<b>K=32</b>	89.28	77.20	/	/	89.10	

Table 3: SupPR Baseline Results. SupPR involves contrastive pretraining of a vision encoder with around 315 million parameters, which contradicts our goal of efficient LVM adaptation. Our method, MVPS, uses a frozen vision transformer and a lightweight prompt retriever that is only 7% the size of SupPR. While SupPR offers valuable insights in the natural image domain, its primary focus is not on optimizing prompt selection for medical imaging. MVPS achieves better gain than SupPR’s which has ten times our training parameters. The significant domain shift in medical images poses challenges for SupPR, which relies on cosine-similarity matching. In contrast, MVPS uses task augmentation during meta-training, resulting in better generalization across domains.