Supplementary Material for "VLSM-Adapter: Finetuning Vision-Language Segmentation Efficiently with Lightweight Blocks"

Table 2. Evaluation of variants of VLSM-Adapters across medical image datasets. **Bold** shows the best score among all the models, and <u>underline</u> represents the superior performance among shallow and dense adapters for the same setting.

| Datasets	Metrics	Adapter Configuration					
		V-Adapter		*VL-Adapter*		*VLC-Adapter*	
		Shallow	Dense	Shallow	Dense	Shallow	Dense
Kvasir-SEG	DSC (%) ↑	86.52	**<u>89.65</u>**	86.85	<u>89.10</u>	86.80	<u>89.06</u>
	IoU (%) ↑	79.05	**<u>82.90</u>**	79.26	<u>82.39</u>	79.33	<u>82.28</u>
	HD95 ↓	53.03	**<u>43.48</u>**	52.18	<u>47.79</u>	53.18	<u>48.82</u>
BKAI	DSC (%) ↑	83.87	<u>86.99</u>	83.67	**<u>87.23</u>**	83.95	<u>87.17</u>
	IoU (%) ↑	75.35	<u>79.60</u>	75.02	**<u>79.81</u>**	75.61	<u>79.74</u>
	HD95 ↓	80.73	**<u>64.48</u>**	87.79	<u>70.02</u>	79.76	<u>64.72</u>
ClinicDB	DSC (%) ↑	<u>89.14</u>	88.95	<u>89.04</u>	88.73	89.04	**<u>89.63</u>**
	IoU (%) ↑	82.02	<u>82.19</u>	<u>81.93</u>	81.84	81.95	**<u>82.89</u>**
	HD95 ↓	<u>18.98</u>	19.12	<u>18.03</u>	18.76	17.77	**<u>17.09</u>**
ISIC-16	DSC (%) ↑	91.59	<u>92.00</u>	91.40	**<u>92.05</u>**	91.55	<u>92.02</u>
	IoU (%) ↑	85.30	<u>85.96</u>	85.05	**<u>85.98</u>**	85.22	<u>85.90</u>
	HD95 ↓	59.56	<u>52.46</u>	60.29	<u>54.38</u>	59.25	**<u>51.76</u>**
DFU	DSC (%) ↑	69.84	<u>71.35</u>	69.47	**<u>72.14</u>**	68.50	<u>71.68</u>
	IoU (%) ↑	58.57	<u>60.59</u>	58.27	**<u>61.42</u>**	57.33	<u>60.69</u>
	HD95 ↓	**<u>38.72</u>**	40.17	<u>38.75</u>	38.79	<u>40.07</u>	41.12
CAMUS	DSC (%) ↑	87.21	<u>89.62</u>	87.16	**<u>89.71</u>**	87.49	<u>89.62</u>
	IoU (%) ↑	78.06	<u>81.70</u>	78.01	**<u>81.85</u>**	78.46	<u>81.71</u>
	HD95 ↓	18.93	<u>14.18</u>	19.14	**<u>14.16</u>**	18.21	**<u>14.16</u>**
BUSI	DSC (%) ↑	64.31	<u>65.05</u>	**<u>65.51</u>**	65.02	<u>64.52</u>	62.95
	IoU (%) ↑	56.82	<u>57.18</u>	**<u>58.19</u>**	57.20	<u>56.88</u>	54.95
	HD95 ↓	69.23	**<u>59.41</u>**	<u>63.36</u>	64.37	<u>65.93</u>	75.99
CheXlocalize	DSC (%) ↑	57.52	<u>58.78</u>	58.14	**<u>58.99</u>**	57.97	<u>58.33</u>
	IoU (%) ↑	44.36	<u>45.74</u>	44.84	**<u>46.01</u>**	44.72	<u>45.33</u>
	HD95 ↓	<u>535.44</u>	536.76	**<u>533.04</u>**	535.97	542.80	<u>533.12</u>
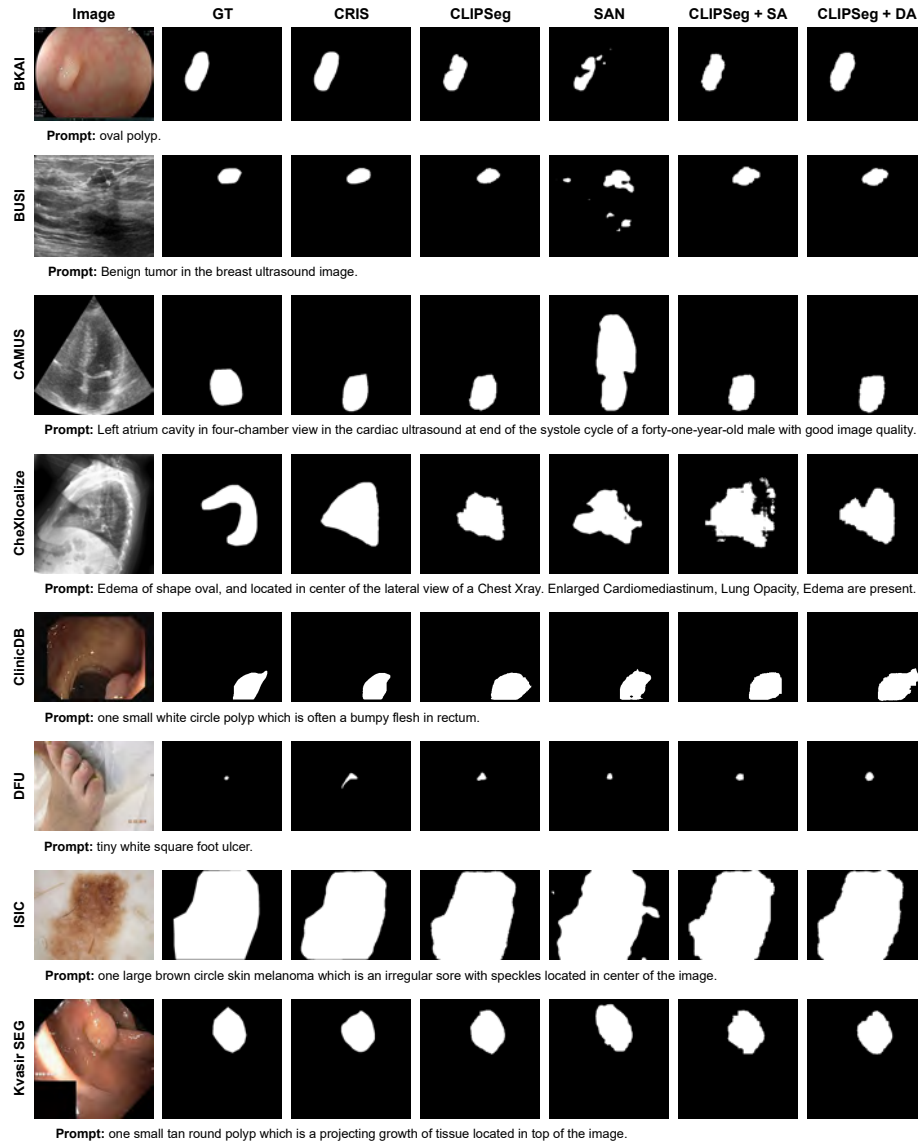
Fig. 4. Qualitative results of different fine-tuned models.