Supplementary Material for Paper #10134

This manuscript is the supplementary material for CVPR 2023 paper #10134, titled *Rethinking Few-Shot Medical Segmentation: A Vector Quantization View*. We further introduce other ablative experimental results of our proposed method on Abdomen MRI dataset [1].

1. Different Number of GFVQ prototypes

Fig. 1 shows the effect of the increasing number of local prototypes generated by GFVQ on few-shot segmentation performance. A threshold of mean performance can be observed, beyond which the performance drops as the number of prototypes increases. This phenomenon is mentioned in the Introduction of paper #10134 and motivated us to develop the better prototype representation mechanism.

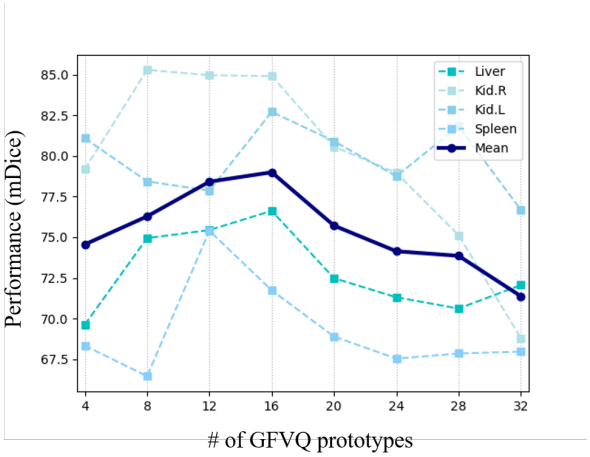

Figure 1. Performance with different number of GFVQ prototypes.

2. Impact of the Number of Iterations

In order to investigate the impact of the max iteration T_s and T_r of SOVQ and ROVQ respectively in an inference, we conduct the extended ablative experiments shown in Tab. 1.

Max iteration T_s of SOVQ is the number of iterative self-organized clustering. As shown in Tab. 1, the larger T_s, the better the performance, which demonstrates the distribution of feature points are better fitted by SOVQ.

#		Liver	Kid.R	Kid.L	Spleen	Mean
T_s	1	75.22	88.88	91.28	67.93	80.83
	10	74.95	88.80	91.42	68.29	80.86
	10^2	75.52	88.05	91.77	68.53	80.97
	10^3	75.17	91.12	90.31	**72.49**	82.27
	10^4	**77.99**	**91.27**	**92.54**	70.44	**83.06**
T_r	1	78.05	91.69	**92.77**	70.96	83.37
	10	77.36	91.53	92.38	73.88	83.79
	10^2	79.92	91.56	89.54	**77.21**	**84.56**
	10^3	81.97	**94.23**	91.96	68.96	84.28
	10^4	**82.51**	93.38	92.20	68.94	84.26

Table 1. Extended ablative results (in Dice score) of different number of max iteration T_s and T_r of GFVQ and SOVQ on Abdomen MRI dataset, respectively.

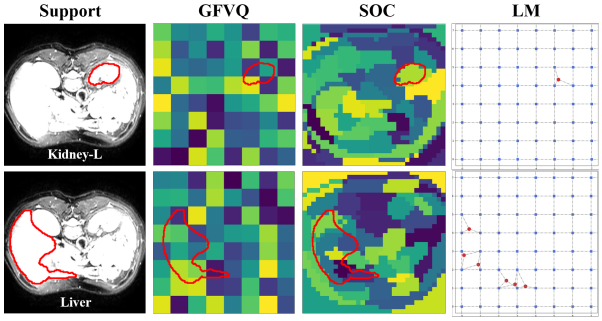

Figure 2. Visualization of SOVQ maps. From left to right: support image, GFVQ clustering map, SOC map and the LM topological map, respectively. The red circles enclose the regions of interest. In clustering maps, the prototypes of different local classes are represented in different colors. In topological maps, for the convenience of observation, only foreground prototypes (red points) are presented.

Max iteration T_s of ROVQ denotes the number of learned prototype fine-tuning. The mean performance reaches the maximum, when $T_r = 10^2$ (Tab. 1). Moderate fine-tuning under the guidance of residual information benefits the segmentation performance. Moreover, the max iteration T in SOVQ and ROVQ also denote the number of sampled feature points taken as a reference. Sampling of feature points provides a global view of the prototype and contributes to both representation and generalization of prototype vectors.

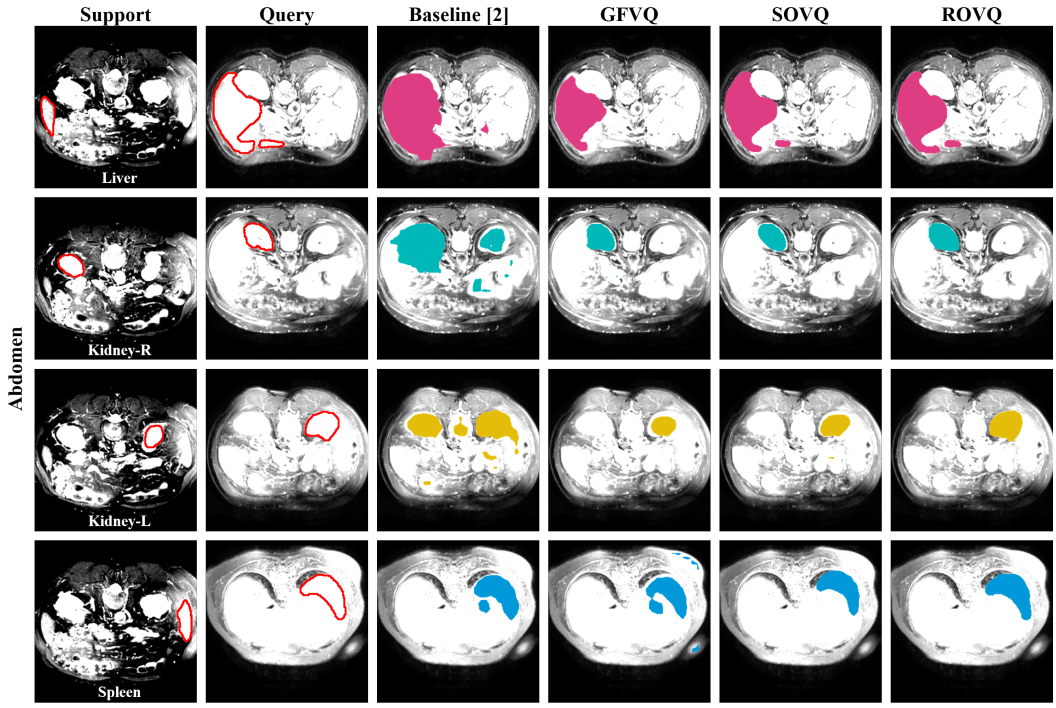

Figure 3. Qualitative results of different components of learning VQ mechanism on Abdomen MRI datasets. GFVQ learns the organ rudimentary; based on GFVQ, SOVQ fleshes out the shape of the foreground and reduces the false alarms; ROVQ further distinguishes the fore- and back-ground, and obtains the more accurate and clear edge.

3. Visualization of the SOVQ map

To demonstrate the effect of the prototype learning of SOVQ, we perform visualization on self-organized clustering (SOC) map and local mapping (LM) topological map, respectively.

SOC map, shown in the Fig. 2, presents the clusters in support image. Compared with the GFVQ clusters, SOC local clusters located in the regions of interest are formed adaptively and outline more accurate object shape.

LM topological map places the SOVQ prototypes in the GFVQ prototype array and form a topological pattern (Fig. 2). The mapped SOVQ prototype reflects the relative relationships with the nearest neighbor GFVQ prototypes and is located in the region of interest as expected.

4. Qualitative results of different components of learning VQ mechanism

Fig. 3 shows the segmentation maps of different components of learning VQ mechanism. The baseline denotes the pipeline in [2] without adaptive local prototype (ALP) learning. As shown in Fig. 3, compared with the baseline, GFVQ initially learns the rudimentary forms of the objects. However, on one hand, since GFVQ excludes most edge prototypes from the foreground, the foreground learned by GFVQ are generally small with blurred edges; on the other hand, some false alarms are also introduced by GFVQ. Then, SOVQ picks up more ignored regions of interest fleshing out the shapes of the foreground and reduces the false alarms. Finally, ROVQ further fine-tunes the edge area and depicts more accurate shape. In summary, with progressive GFVQ, SOVQ and ROVQ, the segmentation results are more closer to the groundtruth masks.

References

[1] A Emre Kavur, N Sinem Gezer, Mustafa Barış, Sinem Aslan, Pierre-Henri Conze, Vladimir Groza, Duc Duy Pham, Soumick Chatterjee, Philipp Ernst, Savaş Özkan, et al. Chaos challenge-combined (ct-mr) healthy abdominal organ segmentation. *Medical Image Analysis*, 69:101950, 2021. 1

[2] Cheng Ouyang, Carlo Biffi, Chen Chen, Turkay Kart, Huaqi Qiu, and Daniel Rueckert. Self-supervision with superpixels: Training few-shot medical image segmentation without annotation. In *European Conference on Computer Vision*, pages 762–780. Springer, 2020. 2