Rethinking Few-Shot Medical Segmentation: A Vector Quantization View

Shiqi Huang, Tingfa Xu*, Ning Shen, Feng Mu, Jianan Li*
Beijing Institute of Technology
{bitsqhuang,lijianan15}@gmail.com, {ciom_xtf1, bitmufeng}@bit.edu.cn, shennbit@163.com

Abstract

The existing few-shot medical segmentation networks share the same practice that the more prototypes, the better performance. This phenomenon can be theoretically interpreted in Vector Quantization (VQ) view: the more prototypes, the more clusters are separated from pixel-wise feature points distributed over the full space. However, as we further think about few-shot segmentation with this perspective, it is found that the clusterization of feature points and the adaptation to unseen tasks have not received enough attention. Motivated by the observation, we propose a learning VQ mechanism consisting of grid-format VQ (GFVQ), self-organized VQ (SOVQ) and residual oriented VQ (ROVQ). To be specific, GFVQ generates the prototype matrix by averaging square grids over the spatial extent, which uniformly quantizes the local details; SOVQ adaptively assigns the feature points to different local classes and creates a new representation space where the learnable local prototypes are updated with a global view; ROVQ introduces residual information to fine-tune the aforementioned learned local prototypes without re-training, which benefits the generalization performance for the irrelevance to the training task. We empirically show that our VQ framework yields the state-of-the-art performance over abdomen, cardiac and prostate MRI datasets and expect this work will provoke a rethink of the current few-shot medical segmentation model design. Our code will soon be publicly available.

1. Introduction

Semantic segmentation is one of the fundamental tasks in medical imaging applications, *e.g.*, disease diagnosis [1, 2], monitoring [3, 4], and screening [5]. With sufficient labeled data being fed into the deep network, segmentation models can achieve promising results. However, in most practical scenarios, the segmentation models often suffer from the lack of required data due to the expensive cost of expertise dense annotations and limited number of abnormal organ and rare lesion samples.

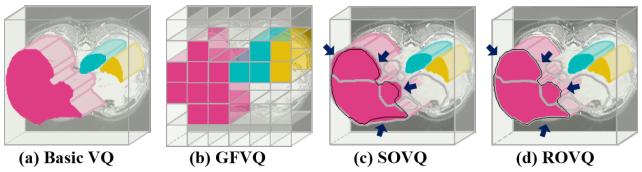

Figure 1. Schematic diagram of different clustering representation schemes of few-shot medical segmentation: (a) the basic VQ with each class represented by a prototype vector; (b) the GFVQ, *i.e.*, the existing local prototype generation, extracts prototype array via mobile pooling window; (c) the proposed SOVQ assigns the pixel-wise features to multiple local classes adaptively; (d) the proposed ROVQ fine-tunes the learned prototype vectors parameterlessly to enhance the adaption performance to unseen tasks. The arrows denote the more accurate edge in (d).

Recently, few-shot medical image segmentation has been widely studied to reduce the requirement for large-scale datasets with dense annotations [6–8]. Currently, the common inference paradigm of few-shot medical image segmentation is to encode a *prototype* to represent the novel class appearing in the support image (Fig. 1(a)) and compute the similarity with query features to perform segmentation [9–12]. The essential work of such a framework lies in prototype learning, which is carried out only by the feature encoder. This encoder is learned with training tasks in the training stage and generalized to unseen tasks in the testing stage. From a *vector quantization (VQ)* view, the prototype vectors representing different classes can be considered as the known sample points in a coding space, and the pixel-wise query feature points are supposed to be classified by decision boundaries determined by the known support points [13–15]. In this view, the prototype learning problem is rethought to be a VQ optimization problem and the prototype vectors learned from support features are thought to serve as the support vectors delineating the encoding space for query features. Therefore, the aim of prototypical few-

*Correspondence to: Tingfa Xu and Jianan Li.

shot segmentation task translates into the requirements for the prototype vectors learned by VQ: discriminative representation and strong generalization.

The requirement for discriminative representation is of concern to many researchers as the prototype generation strategy. Ouyang et al. [10] applied non-overlapping pooling windows to support features generating multiple local prototypes; Yu et al. [11] extracted prototype arrays in the presence of grid constraint and performed a location-guided comparison; Li et al. [12] designed a registration mechanism to align local prototypes between support and query features. The aforementioned schemes can be summarized intuitively that the more prototypes, the better the segmentation performance. However, experiments show that as the number of prototypes increases, the performance deteriorates: on one hand, the set of pooling prototypes reaches saturation of representation capacity; on the other hand, too many prototypes cannot distinguish between classes, resulting in blurred edges or even misclassification. Unlike the requirement for discriminative representation, requirement for strong generalization is often ignored by the previous works in prototype learning. To improve the generalization capability, most researches adopt a unified lightweight encoding network to simultaneously process support and query images [7, 16]. However, few efforts have put generalization studies on prototype learning.

To meet the requirement for discriminative representation, we detail two sub-requirements, *i.e.*, ❶ the clustering of feature points and ❷ the embedding of prototype vectors. Considering this two sub-requirements, we propose a self-organized vector quantization (SOVQ) method, inspired by self-organized mapping algorithm [17, 18], containing a self-organized clustering (SOC) and a local mapping (LM). To abstract features more exactly, SOVQ first creates a new neuron representation space, where neurons are initialized as prototypes and arranged in a normative array. Then the feature points are assigned to different neurons adaptively (for ❶), and the learnable neurons are optimized by the features collaboratively (for ❷). Through iterative learning, the feature points are clustered reasonably and each cluster is represented by a neuron with a global view (Fig. 1(c)).

Furthermore, LM strategy is designed to remap neurons to the encoding space ensuring the prototypes and query features are embedded consistently. Each neuron is interpreted as a weighted sum of GFVQ prototypes via inverse distance weighting and interpolated to GFVQ forming a topologically prototype layout. In summary, through self-organizing the feature points in an unsupervised manner, SOVQ fits the space of interest.

The requirement for strong generalization is also divided into two sub-requirements: ❸ to avoid overfitting to training tasks and ❹ to adapt the model to testing tasks. Thus a residual oriented vector quantization (ROVQ) is put forward, which introduces the residual connection to final vector layout and fine-tunes the learned vectors. On the one hand, the parameter-free learning acts as a regularization term in the training phase to prevent overfitting (for ❸); on the other hand, the residual information with labels guides the prototype vector to get closer to its inherent characteristics and differentiate from other classes (for ❹), which contributes to maintaining details and forming a clearer edge (Fig. 1(d)).

Additionally, following the earlier works on multiple prototype generation, we employ a grid-format vector quantization (GFVQ) to obtain a compressed feature points. As shown in Fig. 1(a), the features are rasterized in grid and compressed by average pooling. Although GFVQ and SOVQ both extract prototypes representing local features, SOVQ is equipped with global receptive field and provides a more specific division of the feature space, while GFVQ is restricted in its grid-format receptive field.

Overall, the medical prototypical few-shot segmentation task is formalized as the vector quantization learning for few-shot class representing. To satisfy the requirement for strong representation and generalization, *i.e.*, sub-requirements ❶-❹, we propose a learning VQ mechanism: a dual structure is employed to integrate GFVQ and SOVQ generating well-representative and limited-quantity prototype vector set, and the former serves as compressed feature reference for LM of SOVQ. Then the prototype set is fine-tuned with ROVQ to maintain the detailed information and enhance generalization capability, and finally the dense prediction is performed by similarity measurement. We show our method achieves the state-of-the-art performance on Abdomen, Cardiac and Prostate MR images with extensive experiments.

2. Related Work

Few-shot Segmentation. To tackle the scarcity of pixel-wise annotations, few-shot segmentation task is introduced to segment unseen classes with the support of a limited amount of labeled data [19]. Dong et al. [20] developed a prototypical episode segmentation network [21], which generates a single prototype vector per class from the support image and then compares it with query features to perform segmentation. The prototypical diagram was later adopted in many further research works [22–26] and adapted to medical images [6–11, 27, 28]. Roy et al. [9] made the first attempt to adopt prototypical learning in segmenting abdominal organs in CT images. Ouyang et al. [10] plugged an adaptive local prototype pooling module into segmentation model to balance the fore- and back-ground classes. Yu et al. [11] extracted multiple prototype vectors and performed a location guided comparison with the query images. Li et al. [12] further added alignment to the pro-

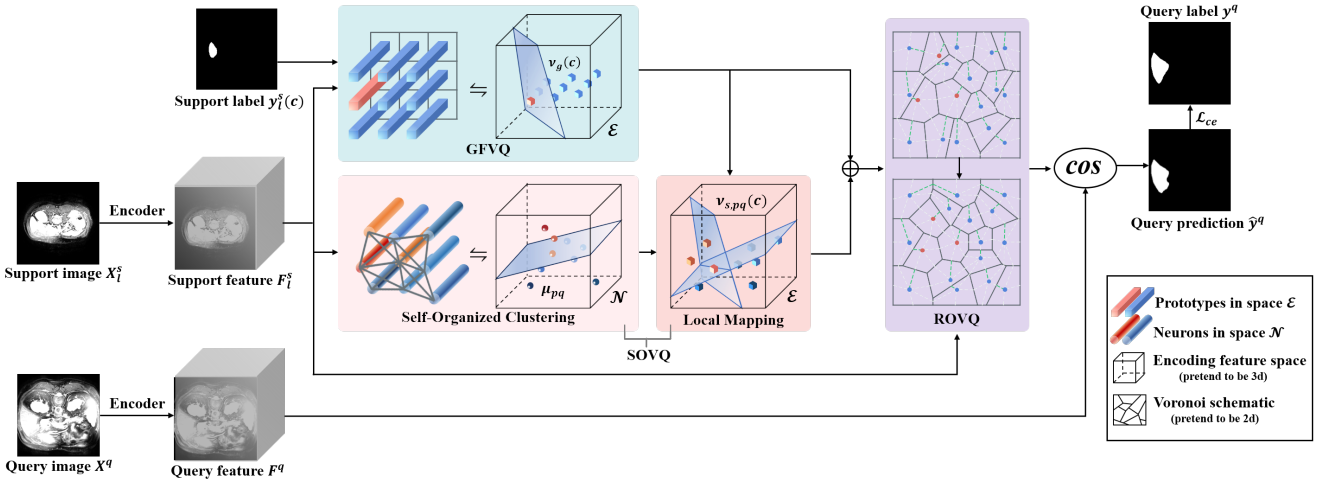

Figure 2. Workflow of the proposed network. The given support image $\mathbf{X}_l^s$ and query image $\mathbf{X}^q$ are first embedded to features $\boldsymbol{F}_l^s$ and $\boldsymbol{F}^q$ in space $\mathcal{E}$ by a shared feature encoder, respectively. Then prototype extraction is performed using the proposed learning VQ mechanism as shown in background color: GFVQ generates the grid-format prototypes $\boldsymbol{\nu}_g(c)$; SOVQ clusters $\boldsymbol{F}_l^s$ adaptively with a new representation space $\mathcal{N}$, where the neurons $\boldsymbol{\mu}_{pq}$ are mapped to $\mathcal{E}$ as $\boldsymbol{\nu}_s(c)$ with the assist of $\boldsymbol{\nu}_g(c)$. both $\boldsymbol{\nu}_g(c)$ and $\boldsymbol{\nu}_s(c)$ are concatenated and fed into ROVQ to fine-tune with residual connection of $\boldsymbol{F}_l^s$. Finally, these prototypes are compared with $\boldsymbol{F}^q$ to output the query prediction map $\hat{y}^q$. For presentation intuition, in GFVQ and SOVQ, the dimension of $\mathcal{E}$ is assumed to be three, so that the vectors are shown as feature points, and in ROVQ it is assumed to be two, in order to show the effect of the voronoi schematic.

totype based on local perception considering the different distribution between the query and support images.

The aforementioned recent studies leveraged lattice pooling strategy for multi-prototype generation, which limits the representational power of these local prototypes. In this work, we propose an adaptive clustering strategy without fixed shape constraints.

Vector Quantization. Vector quantization (VQ) is a concept from signal processing, which models the probability density functions by the distribution of prototype vectors over space [13–15]. In the light of VQ view, the prototypical few-shot segmentation network can be regarded as a combination of clustering and generalization problems that divides the pixel-wise feature points under a constraint of label map. We notice that the above few-shot medical implementations [10–12] argue that more prototypes retain more detailed information. That is, the more clusters portraying, the more accurate the spatial partitioning. However, these prototype vectors derived from the rasterization limit the generalization and representation ability. Self-organized mapping (SOM) algorithm, one of the powerful VQ methods, has attracted our attention [17, 18]. The probability density function is modeled by a set of neurons synaptic weights, which are formed in a typical two-dimensional lattice and is updated collaboratively using unsupervised competitive learning. Recently, SOM is integrated to deep space and achieved the state-of-the-art performance on unsupervised computer vision tasks [29–32].

Inspired by the cortical synaptic plasticity and its self-organization properties of SOM, we incorporate this idea in prototypical few-shot segmentation to enhance the representation and generalization capability. To our best knowledge, it is the first trial of adaptive clustering and unsupervised learning for prototype vectors in few-shot medical segmentation.

3. Existing Prototypical Few-Shot Segmentation

Problem Formulation. The task of few-shot medical segmentation is to allow a model segmenting unseen semantic classes with access to just a few labeled data. In few-shot segmentation, a train set $\mathcal{D}_{tr}$ with training semantic classes $\mathcal{C}_{tr}$, and a test set $\mathcal{D}_{te}$ with testing unseen classes $\mathcal{C}_{te}$, are given, where $\mathcal{C}_{tr} \cap \mathcal{C}_{te} = \emptyset$. The segmentation model is trained on $\mathcal{D}_{tr}$ segmenting semantic classes $\mathcal{C}_{tr}$ and tested on $\mathcal{D}_{te}$ to perform dense prediction on classes $\mathcal{C}_{te}$ without re-training. $\mathcal{D}_{tr} = \{(\mathbf{x}, \mathbf{y}(c)) \mid c = 1, 2, ..., N\}$ is comprised of image $\mathbf{x}$ and its binary segmenting labels $\mathbf{y}(c)$,where N is the number of classes of an episode, and c is the classes from $\mathcal{C}_{tr}$. We define $c = 0$ represents the background class and is neither $\in \mathcal{C}_{tr}$ nor $\in \mathcal{C}_{te}$. $\mathcal{D}_{te}$ is also structured in this way, but with the semantic class set $\mathcal{C}_{te}$.

Prototypical Episode Network. In an episode of few-shot segmentation with prototypical episode network, different images $\mathbf{x} \in$ image space $\mathcal{X}$ and its labels $\mathbf{y}(c) \in$ label space $\mathcal{Y}$ are fed into segmentation model as support set $\mathcal{S}$ and query set $\mathcal{Q}$, respectively. $\mathcal{S} = \{(\mathbf{x}_l^s, \mathbf{y}_l^s(c)) \mid c =$

$0, 1, ..., N; l = 0, 1, ..., K\}$ contains K images $\mathbf{x}^s$ and labels $\mathbf{y}^s(c)$ of N classes, while $\mathcal{Q} = \{\mathbf{x}^q\}$ only contains images $\mathbf{x}^q$ which are supposed be segmented by the knowledge learned from $\mathcal{S}$. The aforementioned segmentation episode is called the N-way K-shot sub-task. In medical image segmentation, most works usually perform 1-way 1-shot learning, which is also adopted in this paper.

Specifically, the first and important step is to generate prototypes $\boldsymbol{p}(c)$, which represent the features of semantic classes. Given $\mathcal{S}$ and $\mathcal{Q}$, a shared feature extractor encodes the support image $\mathbf{x}^s$ and the query image $\mathbf{x}^q$ into support and query features $\boldsymbol{F}^s$ and $\boldsymbol{F}^q$, respectively, where $\boldsymbol{F}^s, \boldsymbol{F}^q \in \mathbb{R}^{H \times W \times D}$ with spatial size (H, W) and embedding dimension D. Then, the prototype $\boldsymbol{p}(c)$ is derived by averaging $\boldsymbol{F}^s$ over the fore- or back- ground regions:

$$\boldsymbol{p}(c) = \begin{cases} f(\boldsymbol{F}_l^s, \mathbf{y}_l^s(c)) & c = 1, 2, ..., N \\ f(\boldsymbol{F}_l^s, 1 - \mathbf{y}_l^s(c)) & c = 0, \end{cases} \tag{1}$$

where

$$f(\boldsymbol{F}_l^s, \mathbf{y}_l^s(c)) = \sum_{x,y}^{H,W} (\boldsymbol{F}_{l,xy}^s \cdot \mathbf{y}_{l,xy}^s(c)) \Big/ \sum_{x,y}^{H,W} \mathbf{y}_{l,xy}^s(c), \tag{2}$$

where x, y represent the coordinates in plane (H, W). Next, we calculate the cosine similarity between each prototype $\boldsymbol{p}(c)$ with $\boldsymbol{F}^q$ to attain the similarity map $\boldsymbol{S}(c)$:

$$\boldsymbol{S}(c) = \frac{\boldsymbol{F}^q \cdot \boldsymbol{p}(c)}{\|\boldsymbol{F}^q\|_2 \|\boldsymbol{p}(c)\|_2}, c = 0, 1, ..., N, \tag{3}$$

where $\cdot$ denotes the dot product between vectors and the $\|\cdot\|_2$ is the second norm function.

Lastly, the similarity maps are assembled as $\{\boldsymbol{S}(c) \mid c = 0, 1, ..., N\}$ and condensed by a softmax function along the class dimension:

$$\hat{\mathbf{y}}^q = \underset{c}{softmax}(\{\boldsymbol{S}(c)\}). \tag{4}$$

where $\hat{\mathbf{y}}^q$ is the final segmentation map.

4. Learning Vector Quantization Few-Shot Segmentation

4.1. Overview.

Generally, previous work employed two prototype vectors, $p(0)$ and $p(\tilde{c})$, to respectively characterize foreground and background features in an 1-way episode. However, when the complete feature encoding space $\mathcal{E}$ is considered, each pixel-wise $\boldsymbol{F}_{l,xy}^s$ is viewed as a data point $\boldsymbol{\xi}_{xy}(c) = (x_1, x_2, ..., x_D)$ in $\mathcal{E}$. Prototype vectors are expected to be the centroids of homogeneous data points. Obviously, only two prototypes on board are not enough to ship the varied object space, *i.e.*, larger quantified and more informative prototype vectors should be introduced to better quantize this space. Further, as a few-shot problem, the generalization ability of prototype is also desired. Motivated by the above analysis, we design a learning vector quantization mechanism containing three components to conduct quantization on feature points $\{\boldsymbol{\xi}_{xy}(c) \mid x = 1, 2, ..., H; y = 1, 2, ..., W; c = 0, 1, ..., N\}$ distributed in encoding space $\mathcal{E}$.

Concretely, at first, the images $\mathbf{x}$ are embedded by a feature encoder: $\mathcal{X} \rightarrow \mathcal{E}$. Then, grid-format vector quantization (GFVQ) is performed on $\mathcal{E}$ to obtain the grid-format prototype vectors through $\mathbf{y}(\tilde{c})$-constrained pooling: $\mathcal{E} \odot \mathcal{Y} \rightarrow \mathcal{E}$. At the same time, self-organized vector quantization (SOVQ) is also performed on $\mathcal{E}$ which adaptively clusters the $\{\boldsymbol{\xi}_{xy}(c)\}$ in a new representation space $\mathcal{N}$ by self-organized clustering (SOC) and remaps the neurons in $\mathcal{N}$ back to $\mathcal{E}$, where segmentation are performed, with local mapping (LM) strategy: : $\mathcal{N}(\mathcal{E}) \rightarrow \mathcal{E}$. After that, the prototype vectors generated by GFVQ and SOVQ are concatenated and are further fine-tuned with the orientation of the residual feature $\boldsymbol{F}^s$ through Residual Oriented Vector Quantization (ROVQ): $\mathcal{E} \rightarrow \mathcal{E}$. Finally, the learned prototype vectors are compared with query features $\boldsymbol{F}^q$ and perform segmentation with the similarity maps: $\mathcal{E} \odot \mathcal{E} \rightarrow \mathcal{Y}$.

4.2. Grid-Format Vector Quantization.

To quantize and compress $\boldsymbol{F}^s$, we first employ GFVQ, a method that uniformly averages local square regions. The same idea has also been adopted in recent works [10–12] and has been proved to be effective in preserving local details. To be specific, $\boldsymbol{F}_l^s$ is first rasterized with grids of size (L_H, L_W), which is the spatial extent to perform average pooling. We define the GFVQ prototype vectors $\boldsymbol{\nu}_g(c)$:

$$\boldsymbol{\nu}_g = \frac{1}{L_H L_W} \sum_{(x,y) \in \Omega} \boldsymbol{F}_{l,xy}^s, \tag{5}$$

where g is index of grid, $g = 1, 2, ..., \frac{HW}{L_H L_W}$, and Ω represents the area of grid g. Then, the corresponding binary label $\mathbf{y}_l^s(\tilde{c})$ with certain foreground class $\tilde{c}$ is also averaged with grid g. The score of foreground pixel ratio δ_g is obtained:

$$\delta_g(\tilde{c}) = \frac{1}{L_H L_W} \sum_{(x,y) \in \Omega} \mathbf{y}_{l,xy}^s(\tilde{c}). \tag{6}$$

The purpose of δ_g is to determine the class of $\boldsymbol{\nu}_g$ by a empirical lower-bound threshold Δ:

$$\boldsymbol{\nu}_g(c) = \begin{cases} \boldsymbol{\nu}_g(c = 0) & \delta_g < \Delta \\ \boldsymbol{\nu}_g(c = \tilde{c}) & \delta_g \geq \Delta. \end{cases} \tag{7}$$

In addition, we prepare the default foreground prototype

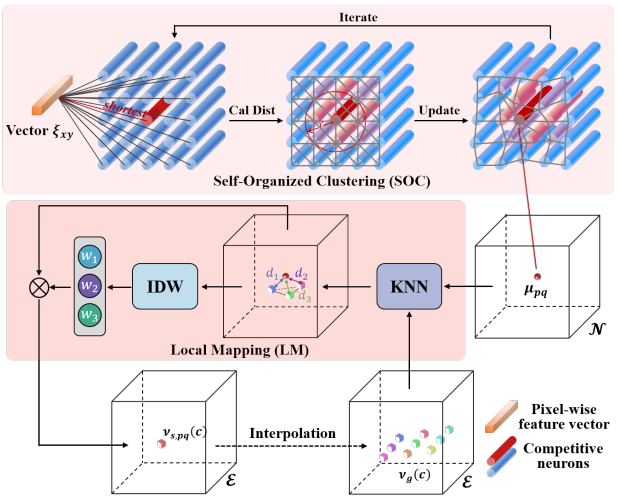

Figure 3. Workflow of the proposed self-organized vector quantization in an iteration, consisting of self-organized clustering and local mapping. The self-organized neuron is learned with pixel-wise feature vector $\boldsymbol{\xi}_{xy}$ randomly selected from $\boldsymbol{F}_l^s$.

vector $\boldsymbol{\nu}'_g(\tilde{c})$:

$$\boldsymbol{\nu}'_g(\tilde{c}) = \sum_{x,y}^{H,W} (\boldsymbol{F}^s_{l,xy} \cdot \mathbf{y}^s_{l,xy}(\tilde{c})) \Big/ \sum_{x,y}^{H,W} \mathbf{y}^s_{l,xy}(\tilde{c}). \quad (8)$$

$\boldsymbol{\nu}_g(c)$ and $\boldsymbol{\nu}'_g(\tilde{c})$ are assembled together as GFVQ prototype vector set $\mathcal{V}_g$. $\mathcal{V}_g$ explicitly compresses the $\boldsymbol{F}_l^s$, which are uniformly divided into multiple clusters and directly characterized by averaging pooling vectors.

4.3. Self-Organized Vector Quantization.

Self-Organized Clustering. Although the $\mathcal{V}_g$ has already realized the local representation in $\mathcal{E}$, the rasterized clustering limits $\mathcal{V}_g$ to adapt to the existing specificity further, especially for the edge regions. In order to address this concern, we design the SOC to enable the adaptive clustering. Inspired by the self-organized mapping algorithm [17, 18], a new representation space $\mathcal{N}$ is created, where $P \times Q$ connected neurons $\{\boldsymbol{\mu}_{pq} \mid p = 1, 2, ..., P; q = 1, 2, ..., Q\} \subseteq \mathcal{N} \subseteq \mathbb{R}^{1\times1\times D}$ are initialized. Through unsupervised learning, $\boldsymbol{\xi}_{xy}(c)$ is assigned to different $\boldsymbol{\mu}$ adaptively and the $\boldsymbol{\mu}$ is updated collaboratively with a global view.

In the t-th iteration, a feature point $\boldsymbol{\xi}^{(t)}_{xy}(c')$ is randomly sampled. $\{\boldsymbol{\mu}_{pq}\}$ are traversed to calculate the distances between each of them and $\boldsymbol{\xi}^{(t)}_{xy}(c')$. The neuron corresponding to the shortest one is selected as the best matching unit (BMU) $\boldsymbol{\mu}_b$:

$$\boldsymbol{\mu}_b = \underset{(p,q)}{argmin} \|\boldsymbol{\xi}^{(t)}_{xy}(c') - \boldsymbol{\mu}^{(t)}_{pq}\|_2, \quad (9)$$

After the $\boldsymbol{\mu}_b$ and its index (p_w, q_w) is acquired, we calculate the Manhattan distance d, which serves as the indicator defining the nearby neurons collection $\{\boldsymbol{\mu}^{(t)}_n\}$ of $\boldsymbol{\mu}_b$:

$$d_{pq} = |p - p_w| + |q - q_w|. \quad (10)$$

Given the neighborhood radius $r^{(t)}$, the neuron $\boldsymbol{\mu}^{(t+1)}_{pq} \in \{\boldsymbol{\mu}^{(t)}_n\}$, of which $d_{pq} \leq r^t$, is updated cooperatively:

$$\boldsymbol{\mu}^{(t+1)}_{pq} = \boldsymbol{\mu}^{(t)}_{pq} + \sigma^{(t)} \times (\boldsymbol{\xi}^{(t)}_{xy}(c') - \boldsymbol{\mu}^{(t)}_{pq}). \quad (11)$$

The decay rate σ and neighborhood radius r are set to decrease linearly:

$$r/\sigma^{(t+1)} = r/\sigma^{(t)} \times (1 - t/T_s). \quad (12)$$

After T_s iterations, $\{\boldsymbol{\xi}_{xy}(c)\}$ is divided into $P \times Q$ clusters, each of which are represent by a prototype neuron $\boldsymbol{\mu}$. To decide the class c of each $\boldsymbol{\mu}$, we count the proportion ε of foreground feature points $\boldsymbol{\xi}_{xy}(c')$ in corresponding clusters and assign class c' to $\boldsymbol{\mu}$, of which $\varepsilon \geq \Delta$.

Local Mapping. To remap neurons from $\mathcal{N}$ to $\mathcal{E}$ where segmentation is operated, we introduce a LM strategy that unifies $\boldsymbol{\mu}$ into the same embemdding format as $\boldsymbol{\nu}_g$. Take a neuron $\boldsymbol{\mu}_{pq}$ as an example, the distances between $\boldsymbol{\mu}_{pq}$ and each of $\mathcal{V}_g$ are calculated and then k $\boldsymbol{\nu}_g$ with the shortest distances are selected. Based on these, we conduct the inverse distance weighting (IDW) function, hence the self-organized prototype vecttor $\boldsymbol{\nu}_s$ is interpreted as

$$\boldsymbol{\nu}_{s,pq} = \sum_{g=1}^{k} \omega_{pqg} \boldsymbol{\nu}_g \Big/ \sum_{g=1}^{k} \omega_{pqg}, \quad (13)$$

where

$$\omega_{pqg} = 1/\|\boldsymbol{\mu}_{pg} - \boldsymbol{\nu}_g\|_2. \quad (14)$$

Overall, each neuron $\boldsymbol{\mu}_{pq}$ is transformed into a weighted sum of $\boldsymbol{\nu}_g$ and turns into $\boldsymbol{\nu}_{s,pq}$. The self-organized prototype vector set $\mathcal{V}_s = \{\boldsymbol{\nu}_{s,pq} \mid s = 1, 2, ..., P \times Q\}$ is assembled with $\mathcal{V}_g$: $\mathcal{V} = \mathcal{V}_g \bigcap \mathcal{V}_s$.

4.4. Residual Oriented Vector Quantization.

Considering the requirement for strong generalization ability of prototype, we introduce ROVQ to modify the prototypes to get closed to the features of interest with no further re-training.

Specifically, ROVQ updates prototypes during a single forward pass with T_r iterations. Each iteration involves randomly selecting a residual feature vector $\boldsymbol{\xi}^{(t)}_{xy}(c')$, comparing it to all prototypes $\mathcal{V}$, and updating the nearest prototype $\boldsymbol{\nu}^{(t)}(c)$ with the feature vector using Eq. 15. This update strategy is based on LVQ [18], which arranges prototypes to create multiple class regions in the encoded feature space.

$$\boldsymbol{\nu}^{(t+1)}(c) = \begin{cases} \boldsymbol{\nu}^{(t)}(c) + \lambda \times (\boldsymbol{\xi}^{(t)}_{xy}(c') - \boldsymbol{\nu}^{(t)}(c)) & c = c' \\ \boldsymbol{\nu}^{(t)}(c) - \lambda \times (\boldsymbol{\xi}^{(t)}_{xy}(c') - \boldsymbol{\nu}^{(t)}(c)) & c \neq c', \end{cases} \quad (15)$$

where λ is the learning rate. The complete $\mathcal{V}$ is finally achieved after T_r times of fine-tuning, of which $\boldsymbol{\nu}$ is later compared with query features $\boldsymbol{F}^q$ via cosine similarity Eq. (3). We apply the softmax function on the similarity maps $\{\boldsymbol{S}(c)\}$ and obtain the output $\hat{\mathbf{y}}^q$ (Eq. (4)).

5. Experiment

5.1. Datasets

To demonstrate the general applicability of our proposed method under different segmentation scenarios, we perform evaluations under three MRI datasets:

Abdomen MRI is from ISBI 2019 Combined Healthy Abdominal Organ Segmentation Challenge (Task 5) [33]. We utilize four of the segmentation classes: *left kidney (Kid.L), right kidney(Kid.R), spleen and liver* and re-sample the data to have the same spacing of 1.25mm $\times$ 1.25mm $\times$ 7.70mm.

Cardiac MRI is from MICCAI 2019 Multi-sequence Cardiac MRI Segmentation Challenge (bSSFP fold) [34]. The label set contains *left-ventricle blood pool (BP), left-ventricle myocardium (MYO) and right-ventricle (RV)*.

Prostate MRI is a data collection of seven prostate studies [35–41], We normalize the data to spacing of 0.75 mm $\times$ 0.75mm $\times$ 2.5mm and assign eight anatomical structures labels to four folds as [12].

To evaluate 2D segmentation performance on 3D volumetric images, we follow the protocol described in [9]: all input images are re-formated as 2D axial and resized to 256 $\times$ 256 pixels. Additionally, we apply random rotation, translation and scaling for data augmentation.

5.2. Implementation Details

We perform 1-way, 1-shot experiments. When training with the real-labels, we take two class for testing and the rest for training; while with the pseudo-labels generated by [10] in self-supervised manner, all classes are tested together. ResNet101 [42] is adopted, which is pre-trained on part of MS-COCO [43] following the SSL practice [10]. The encoder embeds 3 $\times$ 256 $\times$ 256 sized image to 256×32×32 sized feature map. The maximum iteration T_s for SOVQ and T_r for ROVQ is set to 1000 and 10, respectively. Δ is empirically set as 0.95 and r, σ and λ are initialized as 1.0, 0.2 and 0.0001, respectively. We perform 100k iterations in training stage using SGD with a batch size of 1. The learning rate is set to 0.001 with a stepping decay rate of 0.98 per 1000 iterations. We use the Dice score as evaluation metrics.

5.3. Comparison with State-of-the-Arts

Quantitative Results. We compare our method with other eight methods that report the few-shot segmentation results on the Abdomen, Cardiac and Prostate MRI dataset shown in Tab. 1. Superpixel-based self-supervised learning (SSL) is a self-supervised learning strategy designed by [10], which generates superpixel pseudo-labels offline for training and reports the state-of-the-art results on few-shot medical segmentation. With the real-label supervised learning, our method achieves the first-class results that average Dice score 72.44 on Abdomen, 68.02 on Cardiac and 52.25 on Prostate MRI datasets. Considering the regions of interest in Prostate dataset consist of more irregular structures and multiple structures, Prostate dataset is more challenging than the other two, which results in the performance on Prostate dataset lagging behind that of the other two.

Integrated the SSL to our method, our method outperforms the others by a large margin of average Dice score 5.72 on Abdomen, 7.96 on Cardiac and 5.40 on Prostate MRI dataset. Note that the margin on the Cardiac dataset is larger than the others, it is because our method performs better in reducing the false alarms, and the object sizes in Cardiac dataset are generally short, leading to more slices with no foreground class existing than the other two datasets. Importantly, in different segmentation scenarios, our method all yields satisfying results and achieves the state-of-the-art performance, illustrating the robust generalization capability.

Quanlitative Results. Fig. 4 shows the qualitative comparison of segmentation maps between ours and the existing best, *i.e.*, SSL-ALP [10]. Although SSL-ALP has achieved good segmentation results, it still fails to effectively deal with the following cases: no visible boundaries (*e.g.*, Abdomen-Spleen and -Kidney-R) and false alarms (*e.g.*, Cardiac-RV2), which are better handled by ours. Further, compared with SSL-ALP, our method has made a significant improvement in fitting the object shape (*e.g.*, Abdomen-Liver, Cardiac-MYO, Cardiac-RV1 and Prostate-F1-BL) and picking the multiple structures up. (*e.g.*, Prostate-BO).

Fig. 5 visualizes the assignment maps of foreground prototypes. For the same number of prototypes, the responds of SOVQ to the regions of interest are better than GFVQ and is particularly effective in depicting the edges. In addition, it is obvious that the higher the number of prototypes, the more adequate the description of the regions of interest is.

5.4. Ablation Study

Effect of VQ mechanism. Ablative experiments are performed to verify the effectiveness of each VQ component and the results are presented in Tab. 2. Firstly, the effect of GFVQ has been confirmed both in Tab. 2 and in previous works [10–12], which preserves the local details and is subsequently utilized as a compressed feature space in our work. Secondly, SOVQ boosts the performance based on GFVQ. We suggest that self-organized clustering (SOC) further adapts the clustering of features, while each clus-

Method	Adbomen					Cardiac				Prostate				
	Liver	Spleen	Kid.R	Kid.L	Mean	BP	MYO	RV	Mean	Fold1	Fold2	Fold3	Fold4	Mean
SENet [9]	29.02	47.30	47.96	45.78	42.51	58.04	25.18	12.86	32.03	-	-	-	-	-
PANet [16]	50.40	40.58	32.19	30.99	38.53	53.64	35.72	39.52	42.96	-	-	-	-	-
ALPNet [10]	62.35	61.32	60.81	58.83	63.17	73.08	49.53	58.50	42.96	-	-	-	-	-
GCN-DE [7]	49.47	60.63	**83.03**	**76.07**	67.30	-	-	-	-	-	-	-	-	-
RPNet [27]	73.51	69.85	70.00	70.48	79.26	-	-	-	-	-	-	-	-	-
LSLPNet [11]	-	-	-	-	-	-	-	-	-	42.09	29.00	32.49	24.46	32.01
3dCANet [12]	-	-	-	-	-	-	-	-	-	59.36	60.38	45.73	37.60	50.77
Ours	**81.72**	**79.08**	68.94	60.03	72.44	**77.82**	**61.10**	**65.13**	**68.02**	**63.77**	**61.32**	**45.80**	**38.11**	**52.25**
SSL-ALPNet [10]	76.10	72.18	85.18	81.92	78.84	83.99	66.74	79.96	76.90	54.50	46.88	66.38	63.50	57.82
SSL-Ours	**79.92**	**77.21**	**91.56**	**89.54**	**84.56**	**89.68**	**78.27**	**86.64**	**84.86**	**57.12**	**50.23**	**75.12**	**70.31**	**63.22**

Table 1. Quantitative results (in Dice score) on Abdomen, Cardiac and Prostate MRI datasets, respectively.

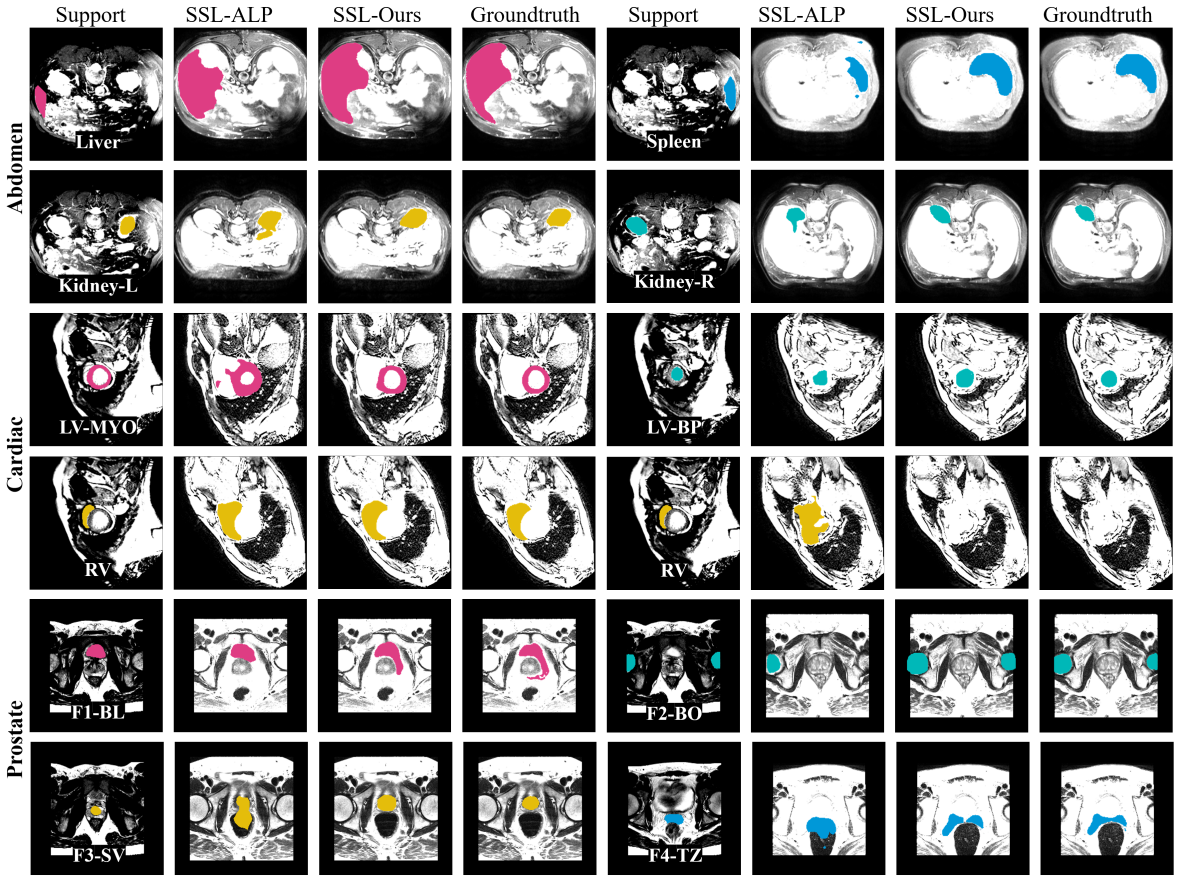

Figure 4. Qualitative results on Abdomen, Cardiac and Prostate MRI datasets. The proposed method achieves satisfying segmentation results which are close to groundtruth. Compared with the previous best-performing SSL-ALPNet, our method accomplishes more outstanding work in fitting object shapes and reducing false alarms.

ter is better represented by the updating neuron instead of pooling prototype. Moreover, the learned neurons are independent of the training task in a unsupervised manner, which benefits the generalization to unseen tasks. Additionally, local mapping (LM) also contributes to the improvement, since the similarity calculation between prototypes and query features is performed per element, hence we are supposed to embed the prototype vector to be consistent with the query features. Lastly, ROVQ also highlights its presence and demonstrates that it is of value to adaptively fine-tune the prototype representation oriented by residual information.

Impact of the number of GFVQ and SOVQ. The number of prototypes, as a key hyper-parameter of VQ, is of interest to us. Specifically, we first investigate the performance with different quantitative ratios of the combination of GFVQ

GFVQ	SOVQ		ROVQ	Liver	Spleen	Kid.R	Kid.L	Mean
	SOC	LM						
				50.41	42.17	31.76	31.24	38.89
✓				75.73	68.32	86.32	84.17	78.63
✓			✓	75.36	70.24	87.54	85.05	79.55
✓	✓			75.38	72.06	84.35	85.53	79.33
✓	✓	✓		78.39	70.58	88.26	86.92	81.04
✓	✓	✓	✓	**79.92**	**77.21**	**91.56**	**89.54**	**84.56**

Table 2. Ablative results (in Dice score) of different components of vector quantization mechanism on Abdomen MRI dataset.

SOVQ share	Liver	Spleen	Kid.R	Kid.L	Mean
0%	68.82	57.34	77.65	67.16	67.74
25%	73.14	59.84	82.89	72.37	72.06
50%	**75.63**	**63.90**	**86.83**	**74.21**	**75.14**
75%	72.01	60.45	82.67	71.11	71.56
100%	61.01	48.09	70.84	59.45	59.85

Table 3. Ablative results (in Dice score) of different shares of SOVQ prototype vector in total 16 prototype vectors on Abdomen MRI dataset. Note the 100% case is free of local mapping.

GFVQ	SOVQ	Liver	Spleen	Kid.R	Kid.L	Mean
4x4	7x7	78.88	75.21	90.07	85.64	82.45
8x8		**79.92**	**77.21**	**91.56**	**89.54**	**84.56**
16x16		78.82	73.28	88.86	84.66	81.41
32x32		77.53	70.88	85.33	84.47	79.55
8x8	3x3	78.20	73.23	87.49	85.81	81.18
	5x5	78.48	75.50	88.48	87.37	82.46
	7x7	**79.92**	**77.21**	**91.56**	**89.54**	**84.56**
	9x9	77.02	75.21	90.12	84.37	81.68

Table 4. Ablative results (in Dice score) of different quantity prototype vector of GFVQ and SOVQ on Abdomen MRI dataset, respectively, where the "$\cdot \times \cdot$" denotes the distribution in 2-D plane.

and SOVQ and secondly with varying numbers in a single variable of GFVQ and SOVQ, respectively.

Tab. 3 compares the effects of different SOVQ shares given a determined total number 16 of prototypes. It can be observed that at the beginning, as the SOVQ share increases, the score also goes up, but after it exceeds 50%, the score starts to go down instead. This may be explained by the fact that although SOVQ prototype shows better representation ability than that of GFVQ, since the SOVQ prototype is expressed by GFVQ, the accuracy of individual SOVQ decreases with the number of GFVQ decreasing.

Tab. 4 shows the performance of the different numbers of GFVQ and SOVQ respectively. As the prototype number increases, the representation of GFVQ or SOVQ tends to saturate and even decreases, while the complementary of the two exhibits a better upper limit of quantity.

Visualization of ROVQ. Ablative experiments of Tab. 2 have demonstrated the effectiveness of ROVQ and we further explore the role taken by ROVQ in the generalization to unseen tasks. Fig. 6 visualizes the similarity between the each training prototype and each testing prototype of certain support and query images respectively. For comparison, the

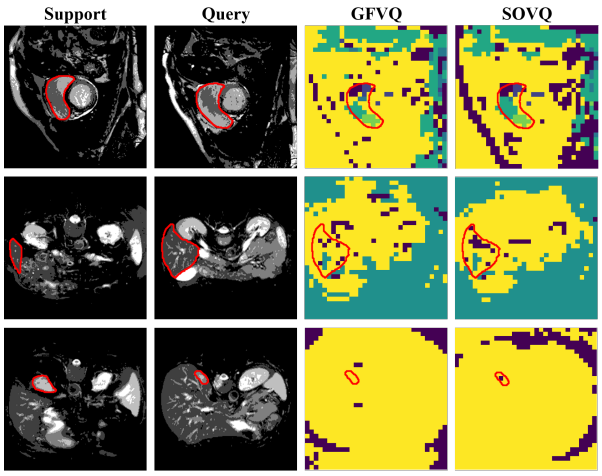

Figure 5. Visualization of foreground prototype assignment maps. The prototypes of different local classes are represented in different colors.

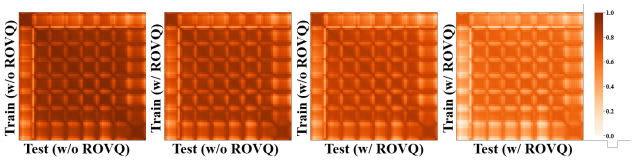

Figure 6. Visualization of pair similarity between training prototypes and testing prototypes on Abdomen dataset.

prototypes are generated directly by adaptive pooling with a kernel of 32×32 and no further class selection is performed. As the involvement degree of ROVQ increases, the overall color of the similarity map gradually becomes lighter, *i.e.*, the similarity decreases. This implies that the adjustment of ROVQ emphasizes the characteristics of the present task at the end of inference, which pulls apart the gap between the prototypes of the training task and the testing task.

6. Conclusion

This work introduces a novel vector quantization view to rethink the prototype learning of few-shot medical segmentation and proposes an effective learning vector quantization mechanism to extract prototype vectors. The aim of the proposed VQ mechanism is to enhance the clustering and representation of prototype and increase the generalization capability to unseen tasks. Our extensive experiments show that the method outperforms the state-of-the-arts on Abdomen, Cardiac and Prostate MRI datasets and confirm the effectiveness of VQ designs. Nevertheless, the self-organizing implementation is hard to handle the complex scenarios. In the future, we will attempt to integrate self-organizing into feature extraction in advance for better unsupervised learning.

References

[1] Dzung L Pham, Chenyang Xu, and Jerry L Prince. A survey of current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(3):315–337, 2000. 1

[2] Neeraj Sharma, Lalit M Aggarwal, et al. Automated medical image segmentation techniques. *Journal of medical physics*, 35(1):3, 2010. 1

[3] Stefan Bauer, Roland Wiest, Lutz-P Nolte, and Mauricio Reyes. A survey of mri-based medical image analysis for brain tumor studies. *Physics in Medicine & Biology*, 58(13):R97, 2013. 1

[4] Matthew J McAuliffe, Francois M Lalonde, Delia McGarry, William Gandler, Karl Csaky, and Benes L Trus. Medical image processing, analysis and visualization in clinical research. In *Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001*, pages 381–386. IEEE, 2001. 1

[5] Feng Zhao and Xianghua Xie. An overview of interactive medical image segmentation. *Annals of the BMVA*, 2013(7):1–22, 2013. 1

[6] Quan Quan, Qingsong Yao, Jun Li, and S Kevin Zhou. Which images to label for few-shot medical landmark detection? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20606–20616, 2022. 1, 2

[7] Liyan Sun, Chenxin Li, Xinghao Ding, Yue Huang, Zhong Chen, Guisheng Wang, Yizhou Yu, and John Paisley. Few-shot medical image segmentation using a global correlation network with discriminative embedding. *Computers in biology and medicine*, 140:105067, 2022. 1, 2, 7

[8] Stine Hansen, Srishti Gautam, Robert Jenssen, and Michael Kampffmeyer. Anomaly detection-inspired few-shot medical image segmentation through self-supervision with supervoxels. *Medical Image Analysis*, 78:102385, 2022. 1, 2

[9] Abhijit Guha Roy, Shayan Siddiqui, Sebastian Pölsterl, Nassir Navab, and Christian Wachinger. 'squeeze & excite'guided few-shot segmentation of volumetric images. *Medical image analysis*, 59:101587, 2020. 1, 2, 6, 7

[10] 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. 1, 2, 3, 4, 6, 7

[11] Qinji Yu, Kang Dang, Nima Tajbakhsh, Demetri Terzopoulos, and Xiaowei Ding. A location-sensitive local prototype network for few-shot medical image segmentation. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 262–266. IEEE, 2021. 1, 2, 3, 4, 6, 7

[12] Yiwen Li, Yunguan Fu, Iani Gayo, Qianye Yang, Zhe Min, Shaheer Saeed, Wen Yan, Yipei Wang, J Alison Noble, Mark Emberton, et al. Prototypical few-shot segmentation for cross-institution male pelvic structures with spatial registration. *arXiv preprint arXiv:2209.05160*, 2022. 1, 2, 3, 4, 6, 7

[13] Nasser M Nasrabadi and Robert A King. Image coding using vector quantization: A review. *IEEE Transactions on communications*, 36(8):957–971, 1988. 1, 3

[14] Allen Gersho and Bhaskar Ramamurthi. Image coding using vector quantization. In *ICASSP'82. IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 7, pages 428–431. IEEE, 1982. 1, 3

[15] Kuilin Chen and Chi-Guhn Lee. Incremental few-shot learning via vector quantization in deep embedded space. In *International Conference on Learning Representations*, 2020. 1, 3

[16] Kaixin Wang, Jun Hao Liew, Yingtian Zou, Daquan Zhou, and Jiashi Feng. Panet: Few-shot image semantic segmentation with prototype alignment. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9197–9206, 2019. 2, 7

[17] Teuvo Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990. 2, 3, 5

[18] Teuvo Kohonen. *Self-organizing maps*, volume 30. Springer Science & Business Media, 2012. 2, 3, 5

[19] Amirreza Shaban, Shray Bansal, Zhen Liu, Irfan Essa, and Byron Boots. One-shot learning for semantic segmentation. *arXiv preprint arXiv:1709.03410*, 2017. 2

[20] Nanqing Dong and Eric P Xing. Few-shot semantic segmentation with prototype learning. In *BMVC*, volume 3, 2018. 2

[21] Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical networks for few-shot learning. *Advances in neural information processing systems*, 30, 2017. 2

[22] Chi Zhang, Guosheng Lin, Fayao Liu, Jiushuang Guo, Qingyao Wu, and Rui Yao. Pyramid graph networks with connection attentions for region-based one-shot semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9587–9595, 2019. 2

[23] Zhuotao Tian, Hengshuang Zhao, Michelle Shu, Zhicheng Yang, Ruiyu Li, and Jiaya Jia. Prior guided feature enrichment network for few-shot segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 2020. 2

[24] Yongfei Liu, Xiangyi Zhang, Songyang Zhang, and Xuming He. Part-aware prototype network for few-shot semantic segmentation. In *European Conference on Computer Vision*, pages 142–158. Springer, 2020. 2

[25] Gen Li, Varun Jampani, Laura Sevilla-Lara, Deqing Sun, Jonghyun Kim, and Joongkyu Kim. Adaptive prototype learning and allocation for few-shot segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8334–8343, 2021. 2

[26] Ye Du, Zehua Fu, Qingjie Liu, and Yunhong Wang. Weakly supervised semantic segmentation by pixel-to-prototype contrast. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4320–4329, June 2022. 2

[27] Hao Tang, Xingwei Liu, Shanlin Sun, Xiangyi Yan, and Xiaohui Xie. Recurrent mask refinement for few-shot medical image segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3918–3928, 2021. 2, 7

[28] Ruiwei Feng, Xiangshang Zheng, Tianxiang Gao, Jintai Chen, Wenzhe Wang, Danny Z Chen, and Jian Wu. Interactive few-shot learning: Limited supervision, better medical image segmentation. *IEEE Transactions on Medical Imaging*, 40(10):2575–2588, 2021. 2

[29] Lyes Khacef, Benoît Miramond, Diego Barrientos, and Andres Upegui. Self-organizing neurons: toward brain-inspired unsupervised learning. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–9. IEEE, 2019. 3

[30] Lyes Khacef, Laurent Rodriguez, and Benoît Miramond. Improving self-organizing maps with unsupervised feature extraction. In *International Conference on Neural Information Processing*, pages 474–486. Springer, 2020. 3

[31] Khodabakhsh Ahmadian and Hamid-Reza Reza-Alikhani. Self-organized maps and high-frequency image detail for mri image enhancement. *IEEE Access*, 9:145662–145682, 2021. 3

[32] Jonas Grande-Barreto and Pilar Gómez-Gil. Pseudo-label-assisted self-organizing maps for brain tissue segmentation in magnetic resonance imaging. *Journal of Digital Imaging*, 35(2):180–192, 2022. 3

[33] 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. 6

[34] Xiahai Zhuang. Multivariate mixture model for myocardial segmentation combining multi-source images. *IEEE transactions on pattern analysis and machine intelligence*, 41(12):2933–2946, 2018. 6

[35] Louise Dickinson, Hashim U Ahmed, AP Kirkham, Clare Allen, Alex Freeman, Julie Barber, Richard G Hindley, Tom Leslie, Chloe Ogden, Rajendra Persad, et al. A multicentre prospective development study evaluating focal therapy using high intensity focused ultrasound for localised prostate cancer: the index study. *Contemporary clinical trials*, 36(1):68–80, 2013. 6

[36] Sami Hamid, Ian A Donaldson, Yipeng Hu, Rachael Rodell, Barbara Villarini, Ester Bonmati, Pamela Tranter, Shonit Punwani, Harbir S Sidhu, Sarah Willis, et al. The smart-target biopsy trial: a prospective, within-person randomised, blinded trial comparing the accuracy of visual-registration and magnetic resonance imaging/ultrasound image-fusion targeted biopsies for prostate cancer risk stratification. *European urology*, 75(5):733–740, 2019. 6

[37] Lucy AM Simmons, Hashim Uddin Ahmed, Caroline M Moore, Shonit Punwani, Alex Freeman, Yipeng Hu, Dean Barratt, Susan C Charman, Jan Van der Meulen, and Mark Emberton. The picture study—prostate imaging (multiparametric mri and prostate histoscanning™) compared to transperineal ultrasound guided biopsy for significant prostate cancer risk evaluation. *Contemporary clinical trials*, 37(1):69–83, 2014. 6

[38] G Litjens, J Futterer, and H Huisman. Data from prostate-3t: The cancer imaging archive, 2015. 6

[39] Geert Litjens, Robert Toth, Wendy van de Ven, Caroline Hoeks, Sjoerd Kerkstra, Bram van Ginneken, Graham Vincent, Gwenael Guillard, Neil Birbeck, Jindang Zhang, et al. Evaluation of prostate segmentation algorithms for mri: the promise12 challenge. *Medical image analysis*, 18(2):359–373, 2014. 6

[40] B Nicolas Bloch, Ashali Jain, and C Carl Jaffe. Data from prostate-diagnosis. the cancer imaging archive. Technical report, accessed 1/18/18, 10.7937, 2015. 6

[41] P Choyke, B Turkbey, P Pinto, M Merino, and B Wood. Data from prostate-mri. *The Cancer Imaging Archive*, 9, 2016. 6

[42] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 6

[43] Hoo-Chang Shin, Holger R Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M Summers. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5):1285–1298, 2016. 6