Appendix

A. Proof of the propositions

Proof of Proposition 1:

Given $\boldsymbol{x}$, let $\boldsymbol{z}^* \in \mathbb{C}$ be the point satisfying $\|\boldsymbol{x}-\boldsymbol{z}^*\|^2 = \min_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2$. Define $w^*(\boldsymbol{z}) := \delta(\boldsymbol{z}-\boldsymbol{z}^*)$ is the dirac delta function. Then we have:

$$\begin{aligned}\min_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 &= \|\boldsymbol{x}-\boldsymbol{z}^*\|^2 \\ &= \sum_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 w^*(\boldsymbol{z}) \\ &\geqslant \min_{w\in\mathbb{W}} \sum_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 w(\boldsymbol{z}).\end{aligned}$$

Moreover, $\forall w \in \mathbb{W}$

$$\sum_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 w(\boldsymbol{z}) \geqslant \|\boldsymbol{x}-\boldsymbol{z}^*\|^2 \sum_{\boldsymbol{z}\in\mathbb{C}} w(\boldsymbol{z}) = \|\boldsymbol{x}-\boldsymbol{z}^*\|^2.$$

so, we can get:

$$\min_{w\in\mathbb{W}} \sum_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 w(\boldsymbol{z}) \geqslant \|\boldsymbol{x}-\boldsymbol{z}^*\|^2.$$

in conclusion,

$$\min_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 = \min_{w\in\mathbb{W}} \sum_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 w(\boldsymbol{z}).$$

Proof of Proposition 2:

Given $\boldsymbol{x}$, let $f(\boldsymbol{z}) := -\|\boldsymbol{x}-\boldsymbol{z}\|^2$, where $\boldsymbol{z}\in\mathbb{C}$. We prove the following equation:

$$\lim_{\varepsilon\to 0^+} \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right) = \max_{\boldsymbol{z}\in\mathbb{C}} f(\boldsymbol{z}).$$

In subsequent derivations, we denote that $f_{max} = \max_{\boldsymbol{z}\in\mathbb{C}} f(\boldsymbol{z})$, then

$$\begin{aligned}\varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right) &= f_{max} + \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})-f_{max}}{\varepsilon}}\right) \\ &\leqslant f_{max} + \varepsilon \ln|\mathbb{C}|.\end{aligned}$$

thus we have $\overline{\lim_{\varepsilon\to 0^+}} \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}} \mathrm{d}\boldsymbol{z}\right) \leq f_{max}$. Notice that

$$\varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right) \geq \varepsilon \ln\left(e^{\frac{f_{max}}{\varepsilon}}\right) = f_{max},$$

we get $\underline{\lim_{\varepsilon\to 0^+}} \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right) \geq f_{max}$. Therefore,

$$\underline{\lim_{\varepsilon\to 0^+}} \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right) = \overline{\lim_{\varepsilon\to 0^+}} \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right) = f_{max}.$$

It implies that

$$\min_{\boldsymbol{z}\in\mathbb{C}} \|x-\boldsymbol{z}\|^2 = -f_{max} = -\lim_{\varepsilon\to 0^+} \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{-\frac{\|x-\boldsymbol{z}\|^2}{\varepsilon}}\right).$$

Proof for the strict concavity of $\min_\varepsilon$ in Proposition 3:

$\forall f_1, f_2 : \Omega \to \mathbb{R}, f_1 \neq f_2$ and $\lambda \in (0,1)$:

$$\min_{\substack{\varepsilon\\ \boldsymbol{z}\in\mathbb{C}}}(\lambda f_1 + (1-\lambda) f_2) = -\varepsilon \ln\left\{\sum_{\boldsymbol{z}\in\mathbb{C}} e^{-\frac{\lambda f_1(\boldsymbol{z}) + (1-\lambda) f_2(\boldsymbol{z})}{\varepsilon}}\right\}.$$

Using the Hölder inequality, we have:

$$\begin{aligned}\sum_{\boldsymbol{z}\in\mathbb{C}} e^{-\frac{\lambda f_1(\boldsymbol{z}) + (1-\lambda) f_2(\boldsymbol{z})}{\varepsilon}} &= \sum_{\boldsymbol{z}\in\mathbb{C}} e^{-\frac{\lambda f_1(\boldsymbol{z})}{\varepsilon}} e^{-\frac{(1-\lambda) f_2(\boldsymbol{z})}{\varepsilon}} \\ &< \left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{-\frac{f_1(\boldsymbol{z})}{\varepsilon}} \mathrm{d}\boldsymbol{z}\right)^{\lambda} \left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{-\frac{f_2(\boldsymbol{z})}{\varepsilon}}\right)^{(1-\lambda)}.\end{aligned}$$

so $\min_{\substack{\varepsilon\\ \boldsymbol{z}\in\mathbb{C}}}(\lambda f_1 + (1-\lambda) f_2) > \lambda \min_{\substack{\varepsilon\\ \boldsymbol{z}\in\mathbb{C}}}(f_1) + (1-\lambda) \min_{\substack{\varepsilon\\ \boldsymbol{z}\in\mathbb{C}}}(f_2)$, which means $\min_\varepsilon$ is strictly concave.

Proof of Proposition 3:

Let $\mathcal{M}_\varepsilon(f) = \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right)$, then

$$\min_\varepsilon(\|\boldsymbol{x}-\boldsymbol{z}\|^2) = -\mathcal{M}_\varepsilon(-\|\boldsymbol{x}-\boldsymbol{z}\|^2).$$

Similar to the above proof of concavity of $\min_\epsilon$, it can get that functional $\mathcal{M}_\varepsilon$ is convex with respect to f. Next, we derive the conjugate representation of $\mathcal{M}_\varepsilon$.

The Fenchel-Legendre transformation of $\mathcal{M}_\varepsilon$:

$$\begin{aligned}\mathcal{M}_\varepsilon^*(w) &:= \max_f \left\{\sum_{\boldsymbol{z}\in\mathbb{C}} f(\boldsymbol{z}) w(\boldsymbol{z}) - \varepsilon \ln\left(\sum_{\boldsymbol{z}\in\mathbb{C}} e^{\frac{f(\boldsymbol{z})}{\varepsilon}}\right)\right\} \\ &= \begin{cases} \varepsilon \sum_{\boldsymbol{z}\in\mathbb{C}} w(\boldsymbol{z}) \ln w(\boldsymbol{z}), & w \in \mathbb{W}, \\ +\infty, & \text{else.}\end{cases}\end{aligned}$$

where $\mathbb{W} = \{w : \mathbb{C} \to [0,1], \sum_{\boldsymbol{z}\in\mathbb{C}} w(\boldsymbol{z}) = 1\}$.

Therefore, we can get the twice Fenchel-Legendre transformation:

$$\mathcal{M}_\varepsilon^{**}(f) = \max_{w\in\mathbb{W}} \left\{\sum_{\boldsymbol{z}\in\mathbb{C}} f(\boldsymbol{z}) w(\boldsymbol{z}) - \varepsilon \sum_{\boldsymbol{z}\in\mathbb{C}} w(\boldsymbol{z}) \ln w(\boldsymbol{z})\right\}.$$

and the above problem has a closed form solution:

$$w_\varepsilon^*(\boldsymbol{z}) = \frac{e^{\frac{f(\boldsymbol{z})}{\varepsilon}}}{\sum_{\boldsymbol{z}'\in\mathbb{C}} e^{\frac{f(\boldsymbol{z}')}{\varepsilon}}}.$$

Since $\mathcal{M}_\varepsilon$ is convex and continuous, it is equal to its twice Fenchel-Legendre transformation.

$$\mathcal{M}_\varepsilon(f) = \max_{w\in\mathbb{W}} \left\{ \sum_{\boldsymbol{z}\in\mathbb{C}} f(\boldsymbol{z})w(\boldsymbol{z}) - \varepsilon \sum_{\boldsymbol{z}\in\mathbb{C}} w(\boldsymbol{z}) \ln w(\boldsymbol{z}) \right\}.$$

then we can get:

$$\begin{aligned}
&\min\nolimits_\varepsilon(\|\boldsymbol{x}-\boldsymbol{z}\|^2) = -\mathcal{M}_\varepsilon(-\|\boldsymbol{x}-\boldsymbol{z}\|^2)\\
&= -\max_{w\in\mathbb{W}} \left\{ \sum_{\boldsymbol{z}\in\mathbb{C}} -\|\boldsymbol{x}-\boldsymbol{z}\|^2 w(\boldsymbol{z}) - \varepsilon \sum_{\boldsymbol{z}\in\mathbb{C}} w(\boldsymbol{z}) \ln w(\boldsymbol{z}) \right\}\\
&= \min_{w\in\mathbb{W}} \left\{ \sum_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2 w(\boldsymbol{z}) + \varepsilon \sum_{\boldsymbol{z}\in\mathbb{C}} w(\boldsymbol{z}) \ln w(\boldsymbol{z}) \right\}.
\end{aligned}$$

the minimizer w_ε^* of the right optimization problem is:

$$w_\varepsilon^*(\boldsymbol{z}) := \frac{e^{\frac{-\|\boldsymbol{x}-\boldsymbol{z}\|^2}{\varepsilon}}}{\sum_{\boldsymbol{z}'\in\mathbb{C}} e^{\frac{-\|\boldsymbol{x}-\boldsymbol{z}'\|^2}{\varepsilon}}}.$$

Proof of Proposition 4:

Let $f_{min} = \min_{\boldsymbol{z}\in\mathbb{C}} \|\boldsymbol{x}-\boldsymbol{z}\|^2$, then

$$\begin{aligned}
\lim_{\varepsilon\to 0^+} w_\varepsilon^*(\boldsymbol{z}) =& \lim_{\varepsilon\to 0^+} \frac{e^{\frac{-\|\boldsymbol{x}-\boldsymbol{z}\|^2}{\varepsilon}}}{\sum_{\boldsymbol{z}'\in\mathbb{C}} e^{\frac{-\|\boldsymbol{x}-\boldsymbol{z}'\|^2}{\varepsilon}}}\\
=& \lim_{\varepsilon\to 0^+} \frac{e^{\frac{f_{min}-\|\boldsymbol{x}-\boldsymbol{z}\|^2}{\varepsilon}}}{\sum_{\boldsymbol{z}'\in\mathbb{C}} e^{\frac{f_{min}-\|\boldsymbol{x}-\boldsymbol{z}'\|^2}{\varepsilon}}}\\
=& \lim_{\varepsilon\to 0^+} \frac{e^{\frac{f_{min}-\|\boldsymbol{x}-\boldsymbol{z}\|^2}{\varepsilon}}}{1+\sum_{\boldsymbol{z}'\in\mathbb{C}\setminus \boldsymbol{c}(\boldsymbol{x})} e^{\frac{f_{min}-\|\boldsymbol{x}-\boldsymbol{z}'\|^2}{\varepsilon}}}\\
=& \delta(\boldsymbol{z}-\boldsymbol{c}(\boldsymbol{x})).
\end{aligned}$$

Thus

$$\begin{aligned}
\lim_{\varepsilon\to 0^+} \boldsymbol{c}_\varepsilon(\boldsymbol{x}) =& \lim_{\varepsilon\to 0^+} \sum_{\boldsymbol{z}\in\mathbb{C}} w_\varepsilon^*(\boldsymbol{z})\boldsymbol{z}\\
=& \sum_{\boldsymbol{z}\in\mathbb{C}} \delta(\boldsymbol{z}-\boldsymbol{c}(\boldsymbol{x}))\boldsymbol{z} = \boldsymbol{c}(\boldsymbol{x}).
\end{aligned}$$

On the other hand, $\lim_{\varepsilon\to+\infty} w_\varepsilon^*(\boldsymbol{z}) = \frac{1}{|\mathbb{C}|}$, then

$$\begin{aligned}
\lim_{\varepsilon\to +\infty} \boldsymbol{c}_\varepsilon(\boldsymbol{x}) =& \lim_{\varepsilon\to +\infty} \sum_{\boldsymbol{z}\in\mathbb{C}} w_\varepsilon^*(\boldsymbol{z})\boldsymbol{z}\\
=& \sum_{\boldsymbol{z}\in\mathbb{C}} \frac{1}{|\mathbb{C}|}\boldsymbol{z} = \hat{\boldsymbol{c}}.
\end{aligned}$$

B. Proof of the theorems

Proof of Theorem 1:

Proof by contradiction. Suppose $\mathbb{S}$ is not a CCS shape with respect to $\partial\mathbb{S}$. According to Definition 3, this implies that there exists a point $\boldsymbol{x}_0 \in \mathbb{S}, \boldsymbol{x}_0 \notin \partial\mathbb{S}$ and a scalar $\lambda_0 \in (0,1)$ such that

$$\boldsymbol{y}_0 = (1-\lambda_0)\boldsymbol{x}_0 + \lambda_0\boldsymbol{c}_\varepsilon(\boldsymbol{x_0}) \notin \mathbb{S}.$$

Let $d_\mathbb{S}(\boldsymbol{x})$ be the Signed Distance Function of set $\mathbb{S}$, defined as follows:

$$d_\mathbb{S}(\boldsymbol{x}) = \begin{cases} -\inf_{\boldsymbol{z}\in\partial\mathbb{S}} \|\boldsymbol{x}-\boldsymbol{z}\|, & \boldsymbol{x}\in\mathbb{S},\\ \inf_{\boldsymbol{z}\in\partial\mathbb{S}} \|\boldsymbol{x}-\boldsymbol{z}\|, & \boldsymbol{x}\notin\mathbb{S}. \end{cases}$$

Now, let $f(\lambda) = d_\mathbb{S}((1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{y}_0)$ for $\lambda\in[0,1]$. Since $d_\mathbb{S}$ is continuous, $f(\lambda)$ is also continuous. Given that $f(0) < 0$ and $f(1) > 0$, there exists some $\lambda_1 \in (0,1)$, such that

$$f(\lambda_1) = 0.$$

Then $\boldsymbol{z}_0 = (1-\lambda_1)\boldsymbol{x}_0 + \lambda_1\boldsymbol{y}_0 \in \partial\mathbb{S}$. We can get:

$$\begin{aligned}
\|\boldsymbol{x}_0 - \boldsymbol{z}_0\|^2 &= \lambda_1^2\|\boldsymbol{x}_0-\boldsymbol{y}_0\|^2\\
&= \lambda_1^2\lambda_0^2\|\boldsymbol{x}_0 - \boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)\|^2\\
&< \|\boldsymbol{x}_0 - \boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)\|^2.
\end{aligned}$$

which contradicts that $\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)$ is a minimizer when $\varepsilon \to 0^+$.

Proof of Theorem 2:

We prove it by contradiction. Suppose $u^{-1}[\gamma,+\infty)$ is not a multi-center star-shape domain.

According to Definition 3, there exists a point $\boldsymbol{x}_0 \in u^{-1}[\gamma,+\infty)$ and a scalar $\lambda_0\in(0,1)$ such that

$$\boldsymbol{y}_0 = (1-\lambda_0)\boldsymbol{x}_0 + \lambda_0\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0) \notin u^{-1}[\gamma,+\infty)$$

which implies that $u(\boldsymbol{y}_0) < \gamma$.

Define $f(\lambda) = u((1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0))$ for $\lambda\in[0,1)$. The function f is continuously differentiable, with the derivative given by

$$f'(\lambda) = \langle\nabla u((1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)), -\boldsymbol{x}_0 + \boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)\rangle.$$

Let $\boldsymbol{y} = (1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)$. Then

$$\begin{aligned}
\boldsymbol{s}(\boldsymbol{y}) &= \boldsymbol{c}_\varepsilon(\boldsymbol{y}) - \boldsymbol{y}\\
&= \boldsymbol{c}_\varepsilon(\boldsymbol{y}) - ((1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0))\\
&= (1-\lambda)(\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0) - \boldsymbol{x}_0) + \boldsymbol{c}_\varepsilon(\boldsymbol{y}) - \boldsymbol{c}_\varepsilon(\boldsymbol{x}_0).
\end{aligned}$$

Based on the definition of $\mathbb{P}$, we have $\langle\nabla u(\boldsymbol{y}), \boldsymbol{s}(\boldsymbol{y})\rangle \geqslant 0$, which means:

$$\langle\nabla u(\boldsymbol{y}), (1-\lambda)(\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0) - \boldsymbol{x}_0) + \boldsymbol{c}_\varepsilon(\boldsymbol{y}) - \boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)\rangle \geqslant 0.$$

In fact, when $\varepsilon\to 0^+$, it is simple to verify $\boldsymbol{c}_\varepsilon(\boldsymbol{y}) = \boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)$. Thus, we can get the following equation when ε is sufficiently small.

$$\langle\nabla u((1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{c}_\varepsilon(\boldsymbol{x}_0)), \boldsymbol{c}_\varepsilon(\boldsymbol{x}_0) - \boldsymbol{x}_0\rangle \geqslant 0.$$

This implies that $f'(\lambda) \geqslant 0$ for all $\lambda\in[0,1)$. Notably, $f'(0)$ corresponds to $f'_+(0)$. Thus, we concluded that $f(\lambda)$ is monotonically increasing for $\lambda\in[0,1)$. Consequently, $f(\lambda_0) \geqslant f(0) \geqslant \gamma$, which contradicts $f(\lambda_0) < \gamma$.

This completes the proof.

C. Details of iterative algorithm Eq. (13)

Once q^{t+1} is obtained, we minimize the following sub-problem w.r.t. u to get u^{t+1}:

$$
\begin{aligned}
u^{t+1} &= \arg\min_{u\in[0,1]}\{\langle -o,u\rangle + \mathcal{H}(u) - \langle q^{t+1}, \boldsymbol{s}\cdot\nabla u\rangle\} \\
&= \arg\min_{u\in[0,1]}\{\langle -o,u\rangle + \mathcal{H}(u) - \langle q^{t+1}\boldsymbol{s}, \nabla u\rangle\} \\
&= \arg\min_{u\in[0,1]}\{\langle -o,u\rangle + \mathcal{H}(u) + \langle div(q^{t+1}\boldsymbol{s}), u\rangle\} \\
&= \arg\min_{u\in[0,1]}\{\langle -o + div(q^{t+1}\boldsymbol{s}), u\rangle + \mathcal{H}(u)\} \\
&= \frac{1}{1+\exp\left(-\frac{o-div(q^{t+1}\boldsymbol{s})}{\xi}\right)}.
\end{aligned}
$$

here the third equality utilizes divergence-gradient duality. The last equality can be obtained by solving it using the following variational method.

Let

$$\mathcal{E}(u) := \min_{u\in[0,1]}\{\langle -o + div(q^{t+1}\boldsymbol{s}), u\rangle + \mathcal{H}(u)\}.$$

Let $u \in [0,1], \forall v \in [0,1]$, construct $u+tv \in [0,1]$, calculate the first-order variation:

$$\frac{\mathrm{d}\mathcal{E}(u+tv)}{\mathrm{d}t}|_{t=0} = \langle v, -o + div(q^{t+1}\boldsymbol{s}) + \xi\ln\left(\frac{u}{1-u}\right)\rangle = 0.$$

utilizing the arbitrariness of the test function v, we can get $u^{t+1} = sigmoid(\frac{o-div(q^{t+1}\boldsymbol{s})}{\xi})$.

From the iterative equations, for an input image of $M\times N$ resolution, the computational complexity of Algorithm 1 is $\mathcal{O}(TMN)$, where T is the number of iterations.

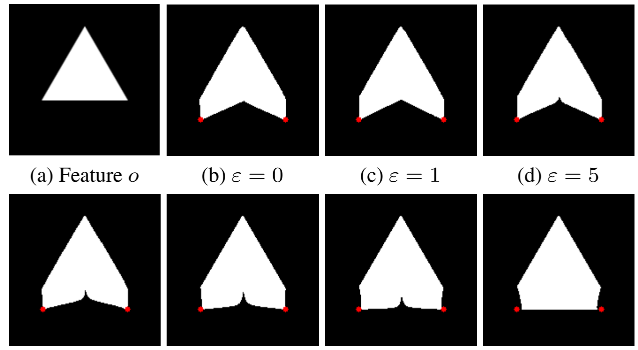
D. Toy experiments

To verify the effectiveness of Algorithm 1, we conducted the toy experiments shown in Fig. 6 and Fig. 7.

For these toy experiments, we set the maximum number of iterations for Algorithm 1 to $T = 3000$ and the coefficient $\xi = 1$. For the experiment in Fig. 7, we set $\varepsilon = 1$, and the segmentation results demonstrate the effectiveness of our algorithm.

E. Hyperparameter design and training strategy

The parameter ε of the CCS shape influences the shape field, which in turn affects the segmentation results, as illustrated in Fig. 6. To obtain the shape field $\boldsymbol{s}$ in the network with CCS module, a token is introduced to learn it, as shown in Fig. 3. Parameters in our CCS network module follow Algorithm 1, with ξ controlling the smoothness of the segmentation indicator function and the dual step size τ_q influencing convergence speed. We found that $\tau_q \le \xi$ is required for algorithm stability. To balance both, we set τ_q and ξ to 1.



(a) Feature o (b) $\varepsilon = 0$ (c) $\varepsilon = 1$ (d) $\varepsilon = 5$

(e) $\varepsilon = 10$ (f) $\varepsilon = 20$ (g) $\varepsilon = 25$ (h) $\varepsilon = 30$

Figure 6. Segmentation results of Algorithm 1 under different smoothing levels of the vector field $\boldsymbol{s} = \boldsymbol{c}_\varepsilon(\boldsymbol{x}) - \boldsymbol{x}$. The red dots represent two points of multi-point star-shaped prior.

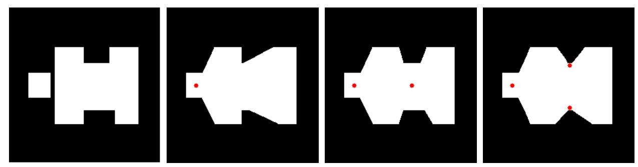


Figure 7. Segmentation results of Algorithm 1 with different center points. The red dots indicate the provided center points.

All experiments in Sec. 4 were conducted using NVIDIA GeForce RTX 4090 GPUs, with each dataset experiment utilizing a single GPU. The batch size was set to 1 for all datasets. The learning rate for the three medical image datasets was uniformly set to 0.0001. The ISIC dataset was trained for 50 epochs, while the Refuge and Kvasir datasets were trained for 30 epochs each. The Adam optimizer was employed for all experiments.

F. More ablations

Ablation Study on the Number of CCS Layers: We also take SAM2 as the backbone to assess the impact of the number of the CCS layers T on WHU Building dataset. The results are shown in Tab. 5, and the optimal value of T is found to be 10 after a comprehensive evaluation.

Table 5. Ablation Study on the Number of CCS Layers on WHU dataset

	WHU		
	DICE↑	*IOU*↑	*ACC*↑
$T = 5$	93.36	88.72	98.62
$\boldsymbol{T = 10}$	93.32	**88.81**	**98.62**
$T = 20$	93.26	88.55	98.60
$T = 30$	**93.39**	88.74	98.62

Optimal Number of Learnable Layers in CCS Module:

We use SAM2 as the backbone to investigate the impact of CCS module with different numbers N of learnable layers. We set the total number of CCS layers T to 10 and conducted ablation experiments on the WHU Building and Refuge datasets. The experimental results are shown in Tab. 6.

Table 6. Ablation Study on the Number of Learnable Layers in CCS Module

Numbers	WHU			Refuge		
	DICE↑	*IOU*↑	*ACC*↑	*DICE*↑	*IOU*↑	*ACC*↑
$N = 10$	93.46	88.86	98.64	81.63	69.89	98.29
$N = 8$	**93.51**	88.89	98.61	**83.76**	**72.82**	**98.65**
$N = 6$	93.47	**88.97**	**98.68**	81.65	70.09	98.30
$N = 4$	93.27	88.69	98.64	81.13	69.39	98.21
$N = 2$	93.35	88.76	98.64	83.21	72.20	98.44

Setting the CCS layer to be learnable can enhance the module's flexibility, enabling it to capture complex features. However, when N is too large, the increased model complexity may lead to overfitting. On the other hand, fixed layers can provide prior information and impose constraints on the solution, helping to prevent overfitting. The results indicate that the performance is optimal when $N = 8$, suggesting that this layer distribution achieves a good balance between flexibility and regularization.

G. Experiments with U-Net[29] as backbone

In this section, we demonstrate the effectiveness of the proposed CCS based on the U-Net [29] framework. We adopt a simple convolutional layer as the sub-network for predicting the vector field. The experimental results are shown in Tab. 7. Fig. 8 provides a visual comparison and demonstration of the experimental results.

Table 7. Result with U-Net[29] as backbone. **Bold** texts stand for the best result.

dataset	model	*DICE*↑	*IOU*↑	*ACC*↑
WHU	U-Net[29]	90.00	84.05	98.10
	CCS loss	90.11	84.16	98.09
	CCS module	**90.38**	**84.44**	**98.10**
ISIC	U-Net[29]	82.64	73.61	96.06
	CCS loss	83.37	74.39	96.17
	CCS module	**84.60**	**76.04**	**96.30**
Refuge	U-Net[29]	60.05	44.84	94.42
	CCS loss	61.74	46.77	95.27
	CCS module	**65.30**	**50.53**	**96.23**
Kvasir	U-Net[29]	81.71	73.65	94.56
	CCS loss	**82.74**	**75.09**	**94.93**
	CCS module	81.94	74.21	94.70

The experimental results demonstrate that our proposed method effectively improves the segmentation performance of U-Net, especially the CCS module, which shows excellent segmentation performance across three datasets.

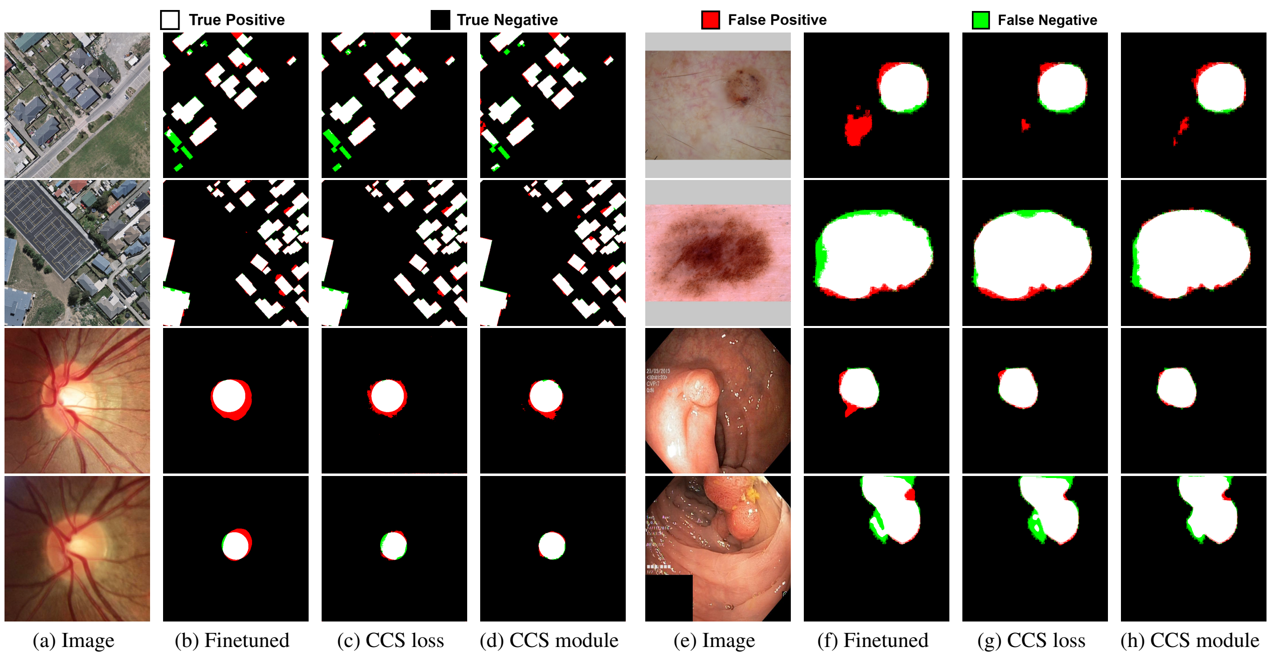


Figure 8. Visual comparison of the experimental results.