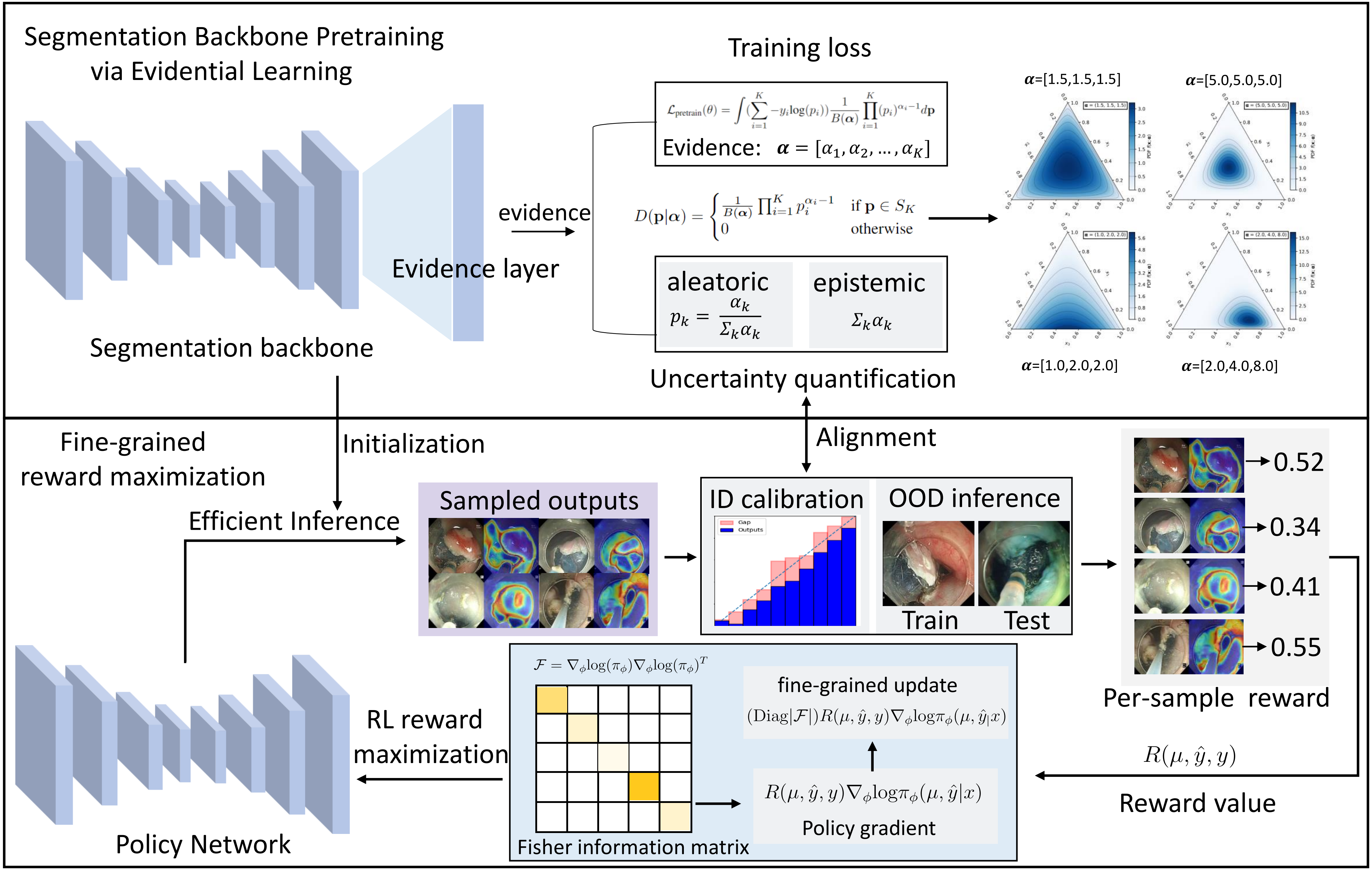

Segmentation Backbone Pretraining via Evidential Learning
Training loss
$\mathcal{L}_{\text{pretrain}}(\theta) = \int (\sum_{i=1}^{K} -y_i \log(p_i)) \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^{K} (p_i)^{\alpha_i - 1} d\mathbf{p}$
Evidence: $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_K]$
evidence
Evidence layer
Segmentation backbone
$D(\mathbf{p}|\boldsymbol{\alpha}) = \begin{cases} \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^{K} p_i^{\alpha_i - 1} & \text{if } \mathbf{p} \in S_K \\ 0 & \text{otherwise} \end{cases}$
aleatoric
$p_k = \frac{\alpha_k}{\Sigma_k \alpha_k}$
epistemic
$\Sigma_k \alpha_k$
Uncertainty quantification
$\boldsymbol{\alpha}$=[1.5,1.5,1.5]
$\boldsymbol{\alpha}$=[5.0,5.0,5.0]
$\boldsymbol{\alpha}$=[1.0,2.0,2.0]
$\boldsymbol{\alpha}$=[2.0,4.0,8.0]
Initialization
Alignment
Fine-grained reward maximization
Efficient Inference
Sampled outputs
ID calibration
Gap
Outputs
OOD inference
Train
Test
0.52
0.34
0.41
0.55
Per-sample reward
$\mathcal{F} = \nabla_\phi \log(\pi_\phi) \nabla_\phi \log(\pi_\phi)^T$
fine-grained update
$(\text{Diag}|\mathcal{F}|) R(\mu, \hat{y}, y) \nabla_\phi \log\pi_\phi(\mu, \hat{y}|x)$
$R(\mu, \hat{y}, y) \nabla_\phi \log\pi_\phi(\mu, \hat{y}|x)$
Policy gradient
Fisher information matrix
RL reward maximization
Policy Network
$R(\mu, \hat{y}, y)$
Reward value