## 6  Supplementary Material

**Alpha scheduling**

$$\alpha = \left( \frac{2}{1 + e^{-10p}} - 1 \right) \cdot \alpha_{\max} \tag{6}$$

with:

$$p = \frac{\text{step} - \text{warmup}_{\alpha}}{\text{total\_steps}} \tag{7}$$

Table 2: Overview of the used datasets, the respective image sizes, and the train, validation, and test splits.

| Dataset | Structures | Patch-Size | Training samples | Validation samples | Test samples |
|---|---|---|---|---|---|
| ACDC | Cardiac | $154 \times 154$ | 4596 | 1110 | 3228 |
| Platelet | (Sub-)Cellular | $200 \times 200$ | 1440 | 360 | 864 |
| OCTA-500 | Retinal | $301 \times 301$ | 128 | 32 | 40 |
| TopCoW | Cerebrovascular | $100 \times 80$ | 70 | 18 | 22 |

Table 3: Hyperparameter search space for model hyperparameters.

| Dataset | max epochs | learning rate samples | num layers | num res. units | batch size |
|---|---|---|---|---|---|
| ACDC | 100 | $[0.0001, 0.01]$ | $\{4, 5\}$ | $\{2, 3, 4, 5\}$ | $\{8, 16, 32, 64, 128\}$ |
| Platelet | 100 | $[0.00005, 0.005]$ | $\{4, 5\}$ | $\{2, 3, 4, 5\}$ | $\{8, 16, 32, 64, 128\}$ |
| OCTA-500 | 300 | $[0.0001, 0.01]$ | $\{4, 5\}$ | $\{2, 3, 4, 5\}$ | $\{8, 16, 32, 64\}$ |
| TopCoW | 300 | $[0.0001, 0.01]$ | $\{4, 5\}$ | $\{2, 3, 4, 5\}$ | $\{8, 16, 32\}$ |

Table 4: Hyperparameter search space for loss-related hyperparameters.

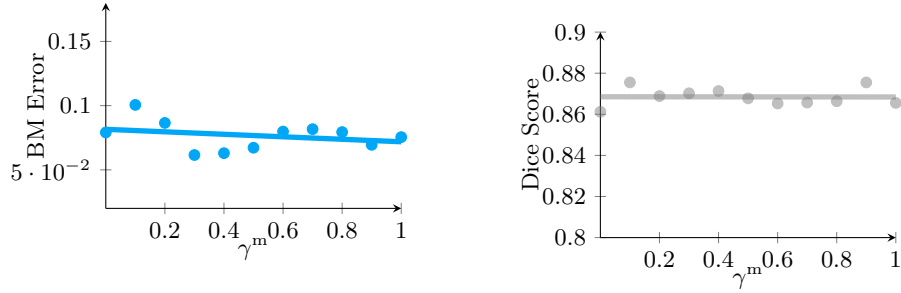| Dataset | $\alpha_{\text{clDice}}$ | $\alpha$ | warmup$_{\alpha}$ | ignore background |
|---|---|---|---|---|
| ACDC | $[0.1, 0.8]$ | $[0.001, 0.1]$ | $\{0, 10, 20, 50\}$ | $\{\text{true}, \text{false}\}$ |
| Platelet | $[0.1, 0.8]$ | $[0.001, 0.1]$ | $\{0, 10, 20, 50\}$ | $\{\text{true}, \text{false}\}$ |
| OCTA-500 | $[0.1, 0.8]$ | $[0.001, 0.05]$ | $\{0, 20, 50, 100\}$ | $\{\text{true}, \text{false}\}$ |
| TopCoW | $[0.1, 0.8]$ | $[0.001, 0.1]$ | $\{0, 10, 20, 50\}$ | $\{\text{true}, \text{false}\}$ |

Fig. 4: Additional ablation on the introduced weighting term with the ACDC dataset. We find a different trend compared to Fig. 3, showcasing that the weight parameter must be tuned according to the dataset.

Table 5: Runtime comparison on two datasets for a single run.

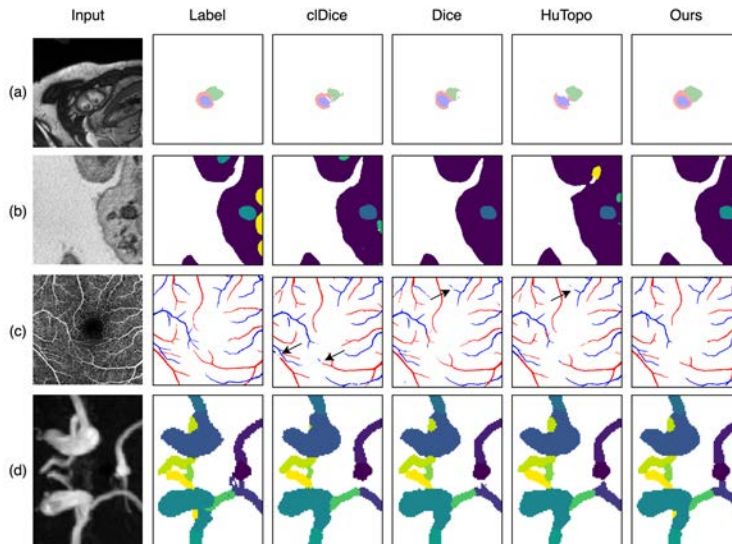| Loss | OCTA-500 | TopCoW |
|------|----------|--------|
| Dice | 16m56s | 5m26s |
| ClDice | 17m17s | 7m32s |
| HuTopo | 59m27s | 19m01s |
| Ours | 28m10s | 10m39s |



Fig. 5: Additional qualitative results on ACDC (a), Platelet (b), OCTA-500 (c), and TopCoW (d) dataset.