Table 1: Hyper-parameters selection for CTTA segmentation on Cityscapes-to-ACDC.

Config	Value
Batch size	1
Optimizer	Adam [**?**]
Weight decay	$(\beta_1, \beta_2) = (0.9, 0.999)$
Warmup epochs	4
Warmup learning rate	8e-5
Learning rate	5e-4
Learning rate schedule	MultiStepLRScheduler ([3, 7], 1e-3)
Augmentation	RandomResizedCrop, RandomHorizontalFlip
Argumentation scale factor	$[0.5, 0.75, 1, 1.25, 1.5, 1.75, 2]$
Ema momentum	0.9995
KL γ	0.1
GCS λ	0.2
Device	[l]RTX3090 $\times$ 1

Table 2: Hyper-parameters selection for Imaganet-to-ImagenetC CTTA.

Config	Value
Batch size	1
Optimizer	SGD
Weight decay	0.9
Learning rate	2e-7
Argumentation scale factor	$[0.5, 0.75, 1, 1.25, 1.5, 1.75, 2]$
KL γ	0.1
GCS λ	0.2
Device	[l]RTX3090 $\times$ 1