

Supplementary Material for: Unsupervised Domain Transfer with Conditional Invertible Neural Networks

Table S1: **Simulated ranges of physiological parameters of a three-layer tissue model for HSI.** $v_{\text{Hb}}[\%]$ represents the blood volume fraction, sO_2 the blood oxygenation, a_{mie} the reduced scattering coefficient at 500 nm, b_{mie} the scattering power, g the scattering anisotropy, n the refractive index, and d the layer thickness. Parameters were uniformly sampled within the specified range.

	$v_{\text{Hb}}[\%]$	$sO_2[\%]$	$a_{mie}[\text{cm}^{-1}]$	$b_{mie}[\text{a.u.}]$	$g[\text{a.u.}]$	$n[\text{a.u.}]$	$d[\text{cm}]$
layers 1 to 3:	0 – 30	0 – 100	5 – 50	0.3 – 3	0.80 – 0.95	1.33 – 1.54	0.002 – 0.2
simulation framework: GPU-MCML, 10^6 photons per simulation							
simulated samples: 5.5×10^5 in wavelength range: 500 nm-1000 nm, step size 2 nm							

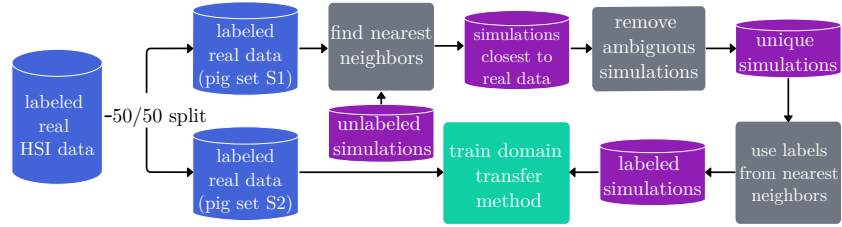


Fig. S1: **HSI data generation.** Real data is leveraged to reduce a large set of simulations to a set of plausible spectra with unambiguous labels.

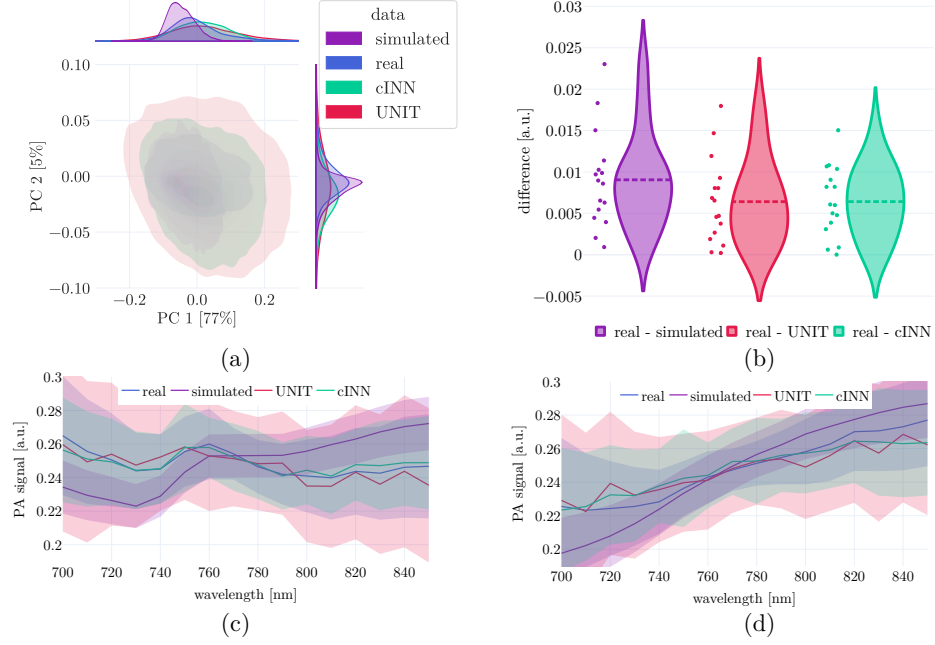


Fig. S2: **Our domain transfer approach yields realistic spectra.** Figures a) and b) are the artery-equivalent of veins in Fig. 5 of the main paper. c) and d) represent the real, simulated, and domain-transferred vein and artery spectra.

Table S2: **Hyperparameters of cINN models and discriminators (Dis) for PAT and HSI.** The hyperparameters have been optimized to yield best spectral consistency and classification performance. The networks have been implemented in PyTorch. Here, “lr” represents the learning rate and “WD” the weight decay. The code and pre-trained models are available at: [GITHUB.LINK](#).

Hyperparameter	PAT cINN	HSI cINN
Epochs	300	300
Batch size	2	10,000
Optimizer	Adam (lr=0.001)	Adam (lr=0.0001)
Optimizer parameters	$\beta_1=0.4, \beta_2=0.999, WD=0.001$	$\beta_1, \beta_2, WD=0.9, 0.95, 0.0001$
Scales	5	1
Blocks per scale	4, 2, 1, 1, 2	40
Conditions per scale	DY, D, D, D, D	30×DY + 10×None
Exponential clamping	1	1
Subnetwork	3×(2D Conv + ReLU) 256 hidden features	2×(Linear + ReLU) 512 hidden features
Hyperparameter	PAT Dis	HSI Dis
Scales	3	1
Layers per scale	4×(2D Conv + LeakyReLU)	3×(Linear + LeakyReLU)
Hidden features	64	256
Dropout prob.	0.2	0.2