A Appendix

	SimCLR	BYOL	MoCov2	SimSiam
Encoder	ResNet-50	ResNet-50	ResNet-50	ResNet-50
Zero init residual	False	False	False	True
Projection model features	MLP (4096, 256)	MLP (4096, 256)	MLP (4096, 256)	MLP (2048, 2048, 2048)
Prediction model features	N/A	MLP (4096, 256)	N/A	MLP (512, 2048)
Momentum encoder (for right encoder)	False	True	True	True
Stop-grad (for right encoder)	False	True	True	True
Contrastive loss temperature	0.1	N/A	0.1	N/A
Optimizer	LARS	LARS	LARS	LARS
Learning rate	0.2	0.2	0.2	0.2
Weight decay	1.5×10^{-6}	1.5×10^{-6}	1.5×10^{-6}	1.5×10^{-6}
Learning rate schedule	cosine decay	cosine decay	cosine decay	cosine decay
Epochs	1000	1000	1000	1000
Linear probe epochs	90	90	90	90
Linear probe learning rate	0.3	0.3	0.3	0.3
Linear probe optimizer	SGD	SGD	SGD	SGD
Linear probe learning rate schedule	cosine decay	cosine decay	cosine decay	cosine decay

Table A.1: **Training setup for each model**: We provide the specific architecture and training setup for each encoder for reproducibility.

Augmentation	Hyper-parameter Values	Probability (Left View)	Probability (Right View)
Random Resized Crop	224×224, min area: 0.08, max area: 1.0, min aspect: 3/4, max aspect: 4. / 3., aspect dist: log, resize method: bicubic	1.0	1.0
Color jitter	contrast: 0.4, brightness: 0.4, saturation: 0.2, hue: 0.1	0.8	0.8
Grayscale	N/A	0.2	0.2
Horizontal flip	N/A	0.5	0.5
Gaussian blur	min sigma: 0.1, max sigma: 2.0, kernel size: 23	1.0	0.1
Amplitude rescale	$m = 0.8, n = 1.75$	0.2	0.0
Phase shift	$p = 0.4, q = 0.7$	0.2	0.0
Random frequency mask	$k \sim [0.01, 0.1)$	0.5	0.0
Gaussian mixture mask	$c = 20, \sigma \sim [10, 15)$	0.2	0.0

Table A.2: **Augmentation hyperparameters:** We provide the parameters used for each augmentation, both image and FDA along with the probability.

A.1 Training Setup

We provide all our implementation details for each baseline - SimCLR, BYOL, MoCov2 and SimSiam in Table A.1. We also include linear probing hyperparameters for full reproducibility.

A.2 Augmentation Hyperparameters

We provide the parameters used for each image and FMA augmentation along with the probability in the left and right view in Table A.2.