This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Supplementary Material

A Hyperparameters

	Pretrain					Finetune/From scratch				
	U-Net		MedNeXt		SwinUNETR	U-Net		MedNeXt		SwinUNETR
Config	XL	B	L	M	-	XL	B	L	M	-
optimizer	AdamW									
base learning rate	$1e-4$					$1e-4$	$1e-4$	$5e-4$	$6e-4$	$1e-4$
weight decay	$3e-5$									
batch size	16	16	8	8	2	4				
learning rate scheduler	cosine decay									
warmup epochs	5	5	5	5	5	0	0	0	0	0
amsgrad	True									
optimizer eps	$1e-8$									
precision	bf16-mixed									
compile	yes	yes	yes	yes	no	yes	yes	yes	yes	no
mask ratio	0.6	0.6	0.6	0.6	0.3	-	-	-	-	-
mask size	4^3				2	-	-	-	-	-
patch size	128^3									

B Augmentations

Table 5: **Augmentations**. p_{sample} denotes the probability a given augmentation is applied to the sample. p_{axis} denotes the probability the given augmentation is applied to each axis, if relevant.

	Augmentation	p_{sample}	p_{axis}	Notes
Spatial	Rotation	0.2	0.66	Rotation in $[-30, 30]$ for all axis
	Scale	0.2	-	Scaling factor in $[0.9, 1.1]$
Intensity	Elastic Deformation	0.33	-	$\alpha \in [200, 600], \sigma \in [20, 30]$
	Gaussian Blur	0.1	-	$\mu = 0, \sigma \in [0, 1]$
	Gaussian Additive Noise	0.2	-	$\mu = 0, \sigma \in [1^{-3}, 1^{-4}]$
	Gaussian Multiplicative Noise	0.2	-	$\mu = 0, \sigma \in [1^{-3}, 1^{-4}]$
	Gamma	0.2	-	Inverts colors of image with $p = 0.01$
	Motion Ghosting	0.2	-	$\alpha \in [0.85, 0.95]$, repetitions $\in [2, 11]$
	Biasfield	0.33	-	-
	Gibbs Ringing	0.2	-	-
	Simulate Low Resolution	0.1	0.33	Zoom range in $[0.5, 1.0]$

C ADNI Acknowledgement

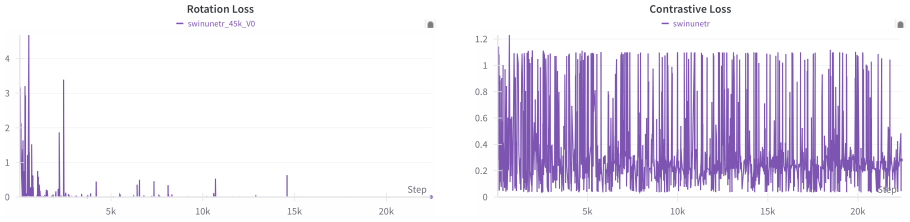
Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimag- ing Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not par- ticipate in analysis or writing of this report. A complete listing of ADNI investigators can be found here.

The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner,MD. The primary goal of ADNI has been to test whether serial magnetic reso- nance imaging (MRI), positron

emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Depart- ment of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.;Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & John- son Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.;Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transi- tion Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research In- stitute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

D Contrastive and Rotation Loss Curves



(a) Rotation loss (b) Contrastive Loss

Fig. 3: Rotation and contrastive losses when training SwinUNETR [23] on 🧠BRAINS-45K.

E Results with standard deviation

Table 6: Main results (as Table 4) with variation at one standard deviation.

	Pretraining Data		**Params** *Millions*	**BraTS21** $n_{\text{test}} = 313$	**ISLES22** $n_{\text{cv}} = 167$	**WMH** $n_{\text{cv}} = 40$	**WMH** OOD $n_{\text{test}} = 110$
U-Net XL	-	From Scratch	90	56.00 ± 2.55	48.30 ± 2.94	68.32 ± 1.75	67.22 ± 1.59
U-Net B	-	From Scratch	22	55.28 ± 1.75	45.66 ± 2.40	66.18 ± 2.28	64.96 ± 2.52
MedNeXt M U-Net dec	-	From Scratch	21	58.38 ± 1.44	56.15 ± 2.77	70.01 ± 1.48	67.64 ± 1.74
MedNeXt L U-Net dec	-	From Scratch	55	58.28 ± 1.23	55.40 ± 1.88	70.32 ± 1.77	67.57 ± 1.48
MedNeXt L MedNeXt dec	-	From Scratch	62	55.42 ± 2.25	53.88 ± 1.52	68.75 ± 2.18	67.14 ± 1.28
SwinUNETR	-	From Scratch	62	50.00 ± 2.59	42.55 ± 2.49	64.96 ± 3.90	65.02 ± 2.86
U-Net XL	🧠-45K	AMAES	90	58.98 ± 0.88	53.58 ± 0.76	70.40 ± 1.46	68.22 ± 1.16
U-Net B	🧠-45K	AMAES	22	57.20 ± 1.75	48.42 ± 1.32	69.11 ± 2.28	67.51 ± 1.56
MedNeXt M	🧠-45K	AMAES	21	56.81 ± 2.40	55.89 ± 2.17	71.70 ± 1.62	68.80 ± 0.79
MedNeXt L	🧠-45K	AMAES	55	58.73 ± 1.33	56.83 ± 1.39	71.63 ± 1.67	68.68 ± 0.97
SwinUNETR	🧠-45K	[23]	62	49.54 ± 0.94	48.42 ± 3.14	67.28 ± 2.67	68.14 ± 1.43