9. Supplementary Material

9.1. Participant and Scan Characteristics

Study	N	Age Mean(SD)	Sex Male/Female	Scanner (Field Strength)	Scan Sequence	Voxel Dimensions
OASIS (Open Access Series of imaging studies)	288	44.1 (23)	106/188	Siemens Vision (1.5T)	MPRAGE	$1.0 \times 1.0 \times 1.25$ mm
IXI (Information eXtraction from Images)	561	48.6 (16.5)	250/311	Philips Intera(3T); Phillips Gyeroscan Intera (1.5T;GE Signa (1.5T);	T1-FFE; MPRAGE	$0.94 \times 0.94 \times 1.2$ mm
ICBM (International Consortium for Brain Mapping)	322	24.8 (5.1)	177/145	Siemens Magnetom (1.5T)	MPRAGE	$1.0 \times 1.0 \times 1.0$ mm
ABIDE (Autism Brain Imaging Data Exchange)	184	16.9 (6.7)	161/23	Various (all 3T)	MPRAGE	Various
Beijing Normal University	179	21.3 (1.9)	72/107	Siemens(3T)	MPRAGE	$1.33 \times 1.0 \times 1.0$mm
Nathan Kline Institute (NKI)/ Rockland	160	41.5 (18.1)	96/64	Siemens Tim Trio (3T)	MPRAGE	$1.0 \times 1.0 \times 1.0$mm
MCIC (Mind Clinical Imaging Consortium)	93	32.5 (12)	64/29	Siemens Sonata/Trio (1.5/3T); GE Signa (1.5T)	MPRAGE; SPGR	$0.63 \times 0.63 \times 1.5$mm
Berlin School of Brain & Mind	49	31 (7.1)	24/25	Siemens Tim Trio (3T)	MPRAGE	$1.0 \times 1.0 \times 1.0$mm
NEO2012	39	29.6 (8.4)	18/21	Siemens Allegra(3T)	MPRAGE	$1.0 \times 1.0 \times 1.0$mm
TRAIN-39	36	22.7 (2.6)	11/25	Siemens Allegra (3T)	MPRAGE	$1.33 \times 1.33 \times 1.3$mm
Cleveland Clinic	31	43.6 (11.1)	11/20	Siemens Tim Trio (3T)	MPRAGE	$1.0 \times 1.0 \times 1.2$mm
WUSL	24	23 (1.4)	4/20	Siemens Tim Trio (3T)	MPRAGE	$1.0 \times 1.0 \times 1.0$mm
MIRIAD (Minimal Interval Resonance Imaging in Alzheimer's Disease)	23	69.7 (7.2)	12/11	GE Signa (1.5T)	3D IR-FSPGR	$0.94 \times 0.94 \times 1.5$mm
CADDementia	12	62.3 (6.3)	9/3	GE Signa (3T)	3D IR-FSPGR	$0.9 \times 0.9 \times 1.0$mm
Dataset Total	2001	37 (18.1)	1016/985	-	-	-

Table 3: Additional scan and participant information of the Brain Age Healthy Cohort dataset [40].

				Vascular		Territory	
Scanner Brand	**Field Strength**	**Resolution**	**N**	*MCA*	*ACA*	*Lacunar*	Other
GE 750 Discovery	3T	1.0 x 1.0 x 1.0 mm	76	12	0	46	18
GE Signa	1.5 T	0.9 x 0.9 x 0.9 mm	31	14	2	11	4
GE Signa Excite	3T	1.0 x 1.0 x 1.0 mm	34	10	0	18	6
GE Signa HD-X	3T	1.0 x 1.0 x 1.0 mm	26	1	0	24	1
Phillips Achieva	3T	1.0 x 1.0 x 1.0 mm	36	15	0	18	3
Siemens Trio	3T	1.0 x 1.0 x 1.0 mm	101	56	1	32	12

Table 4: Additional information about the patients scans of the stroke lesion Data-set (ATLAS R1.2).

9.2. Equal Domain Source Task Hyper-parameters

The CNNs were pre-trained for 30 epochs using a batch size of 32 and batch normalization [52] before the activation function. The encoder and decoder used a learning rate of 10^{-4}, a weight decay rate of $2 \cdot 10^{-6}$ and an Adam optimizer [66]. If a discriminator was used, it was turned on after 26000 iterations. The discriminator used a learning rate of 10^{-5} and an SGD optimizer. The encoder and a fully connected block used the same hyper-parameters as the encoder and decoder.

To ensure that the CNNs pre-trained equal and unequal source domain used a similar amount of data, 870 scans from the BAHC dataset were randomly selected. These scans were randomly split into 670 training scans, 100 validation scans and 100 testing scans.

9.3. Unequal domain Source Task Hyper-parameters

The CNNs were pre-trained as described in the Taskonomy study [15]. The CNNs were pre-trained for 30 epochs using a batch size of 32 and batch normalization [52] before the activation function. The encoder and decoder used a learning rate of 10^{-4}, a weight decay rate of $2 \cdot 10^{-6}$ and an Adam optimizer [66]. If a discriminator was used, it was turned on after 25000 iterations. The discriminator used a learning rate of 10^{-5} and an SGD optimizer. The learning rate was annealed by a factor of 10 after 80000 iterations.

The CNNs were trained on 120000, validated on 16000 and tested on 17000 images.