Brain-ID: Learning Contrast-agnostic Anatomical Representations for Brain Imaging

Peirong Liu[1] Oula Puonti[1] Xiaoling Hu[1]
Daniel C. Alexander[2] Juan E. Iglesias[1,2,3]

[1]Harvard Medical School and Massachusetts General Hospital
[2]University College London [3]Massachusetts Institute of Technology

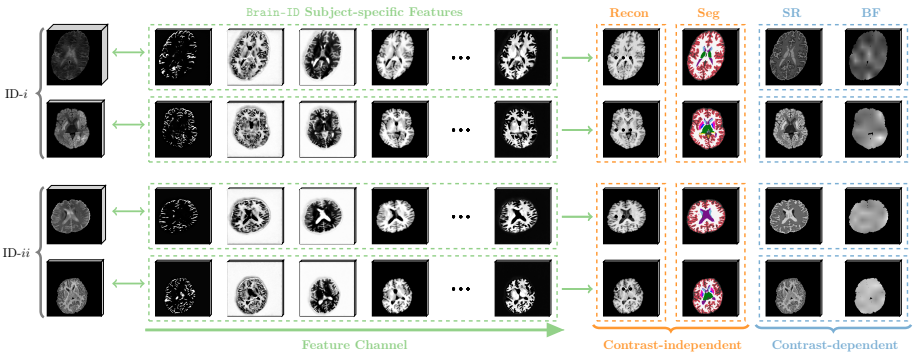

Fig. 1: Brain-ID features serve as the distinctive identity of each subject (ID) regardless of appearance (contrast, deformation, corruptions, etc) and quickly adapt to downstream tasks – either contrast-independent (anatomy **Recon**struction, **Seg**mentation), or contrast-dependent (**S**uper-**R**esolution, **B**ias **F**ield estimation), all through one layer.

Abstract. Recent learning-based approaches have made astonishing advances in calibrated medical imaging like computerized tomography (CT). Yet, they struggle to generalize in uncalibrated modalities – notably magnetic resonance (MR) imaging, where performance is highly sensitive to the differences in MR contrast, resolution, and orientation. This prevents broad applicability to diverse real-world clinical protocols. We introduce Brain-ID, an anatomical representation learning model for brain imaging. With the proposed "mild-to-severe" intra-subject generation, Brain-ID is robust to the subject-specific brain anatomy regardless of the appearance of acquired images. Trained entirely on synthetic inputs, Brain-ID readily adapts to various downstream tasks through one layer. We present new metrics to validate the intra/inter-subject robustness of Brain-ID features, and evaluate their performance on four downstream applications, covering contrast-independent (anatomy reconstruction, brain segmentation), and contrast-dependent (super-resolution, bias field estimation) tasks (Fig. 1). Extensive experiments on six public datasets demonstrate that Brain-ID achieves state-of-the-art performance in all tasks on different MR contrasts and CT, and more importantly, preserves its performance on low-resolution and small datasets. Code is available at https://github.com/peirong26/Brain-ID.

1 Introduction

Magnetic resonance imaging (MRI) enables in vivo noninvasive imaging of the human brain with exquisite and tunable soft-tissue contrast [8]. Recent machine learning methods have achieved great improvements in faster and more accurate MRI analysis [27], such as segmentation [18,32,42,47], registration [3,49,55,64], super-resolution [52,53], and connectivity studies [45]. However, most existing MRI analysis methods are specific to certain MR contrast(s) and require near-isotropic acquisitions. Therefore, models face sharp performance drops when voxel size and anisotropy increase, or are being used for a different contrast than training [57]. This reduces model generalizability and results in duplicate training efforts. Resorting to synthetic data, recent contrast-agnostic models [4, 25, 27, 28, 36, 39, 40] achieve impressive results and largely extend model applicability. However, these models are only applicable to the tasks they were trained for.

Meanwhile, task-agnostic foundation models in computer vision and natural language processing have witnessed remarkable success, along with the fast developments of large-scale datasets [2,7,10,14,34]. However, due to different acquisition/processing protocols and privacy requirements across institutions, large-scale medical imaging datasets require significantly more effort than natural imaging/language. Thus, medical foundation models are not as well developed. MONAI [11] includes a pre-trained model zoo, yet all are highly task-oriented and sensitive to contrasts. Zhou et al. [66] constructed a medical foundation model, designed for eye and systemic health condition detections from retinal scans, and only works on the modalities of color fundus photography and optical coherence tomography. Lately, generalist biomedical AI systems [44, 50, 54] have shown great potential in biomedical tasks under vision-language context, e.g., (visual) question answering, image classification, radiology report generation and summarization. However, they have not explored the challenging vision tasks such as reconstruction, segmentation, super-resolution, and registration.

We introduce `Brain-ID`, a contrast-agnostic feature representation approach trained on synthetic inputs, that can adapt to various tasks through one layer.

1) We introduce an on-the-fly, intra-subject data generator capable of sythesizing *any* contrast, with a *mild-to-severe* corruption strategy (Fig. 2, Sec. 3.2). Unlike real-world datasets constrained by the images acquired per subject, `Brain-ID` learns in a much more expansive and diverse space.
2) We design a feature representation learning framework guided by the *unique anatomy* of each subject. We validate that `Brain-ID`'s high-resolution features are consistently robust to the superficial perturbations of an image's appearance (Sec. 4.2). Straightforward one-layer adaptions are further presented to adapt `Brain-ID` features to downstream tasks (Sec. 3.3).
3) We evaluate `Brain-ID` on both contrast-independent (anatomy reconstruction, brain segmentation) and contrast-dependent (super-resolution, bias field estimation) tasks, across six public datasets ($\approx$ 8,000 images) including MR (T1w, T2w, FLAIR) and CT. `Brain-ID` achieves state-of-the-art performance on all tasks (Tab. 2), and further maintains its high performance on low-resolution images (Fig. 6) and limited-size datasets (Fig. 7).

2 Related work

Feature Representation in Medical Imaging As mentioned in Sec. 1, general feature representation learning in medical imaging can be more challenging than in the natural image domain, due to limited data availability. Xu et al. [62] introduced a multiple-instance learning model for feature representation, but it only applies to histopathology image classification. You et al. [65] presented CVRL, a semi-supervised representations approach designed for image segmentation, but it requires CT to extract anatomical information. SAM [63] encodes anatomical information in feature embeddings, which has shown to be effective for registration [38]. However, same as CVRL, SAM only works on CT. BrainPrint [56] is a compact and discriminative representation of brain morphology, which is specifically designed for cortical surface analyses. To our best knowledge, CIFL [15] is the only existing work on learning contrast-agnostic and task-independent brain feature representations. CIFL relies on contrastive learning alone, and is insufficient to outperform task/contrast-specific supervised models in downstream applications as shown by our experiments (Tab. 2).

Contrast-invariant Learning for MR Images MRI scans acquired across sites vary substantially in contrasts, resolutions, orientations, etc. When given a new dataset, heterogeneity leads to duplicate training efforts for approaches that are sensitive to specific combinations of MR contrast, resolution, or orientation. Classical brain segmentation models used Bayesian inference for contrast-robustness [20, 37], which requires a long processing time and struggles with low or anisotropic resolutions [27, 46]. Recently, there have been works using synthetic data to achieve contrast-invariance in tasks like image segmentation [4,36], registration [25], super-resolution [28], and skull-stripping [26]. However, all the above-mentioned methods are trained in a task-specific manner, their features therefore cannot be readily applied to other domains.

3 `Brain-ID`: Learning Anatomy-specific Brain Features

As discussed in Sec. 1, the main challenges to obtain a general and robust feature representation for MR images lie in *(i)* the practical restrictions of building large-scale datasets with diverse contrasts; and *(ii)* the nature of most medical imaging models that are task-oriented and specific to data type (contrast, resolution, orientation, etc). We aim to learn a brain feature representation that is:
(i) Robust: features should be robust to each subject's distinct anatomy, unaffected by variations in poses/deformations, contrasts, resolutions, or artifacts.
(ii) Expressive: features should also exhibit high expressiveness, containing rich information that facilitates easy and effective adaptation to diverse downstream tasks, eliminating the necessity for extensive training data.

We first introduce `Brain-ID`'s data generator (Sec. 3.1) and training framework (Sec. 3.2) to achieve the above two aims. Then, we present our one-layer solutions for adapting `Brain-ID` features to downstream tasks that could be either dependent or independent of the contrast of input images (Sec. 3.3).

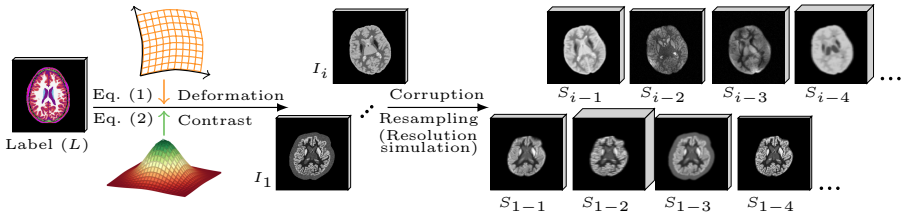

Fig. 2: `Brain-ID`'s data generator on the fly. Given the brain segmentation labels of a subject, we randomly generate a deformation field, and synthesize intra-subjecct samples featuring various contrast intensities and corruption levels (Sec. 3.1).

3.1 Enrich the Intra-subject Learning Space

A good representation relies on large-scale data, however, different acquisition protocols, processing pipelines, and privacy requirements across institutions make large-scale data less accessible and require significantly more effort [1, 17, 48]. Moreover, for each subject, there are usually only limited acquired scans available. This lack of data consistency is a significant barrier to obtaining a robust, contrast-invariant feature representation. `Brain-ID` avoids these barriers through the use of synthetic data.

To generate images with complex brain structures, we start with brain segmentation labels of anatomical structures (L in Fig. 2). The generatation consists of three steps, *(i)* deformation generation, *(ii)* contrast synthesis, and *(iii)* data corruption (including lower resolution resampling). For simplicity, Θ denotes the parameter group of the generation process described below.

Deformation Generation We first generate a random deformation field ($\phi|_{\theta_\phi}$) consisting of an affine transformation and a non-linear displacement field:

$$\phi|_{\theta_\phi} = \mathcal{T}|_{\theta_\phi} \circ \mathcal{A}|_{\theta_\phi}\,, \tag{1}$$

where $\mathcal{A}|_{\theta_\phi}$ denotes an affine transformation matrix which includes linear (rotation, scaling, shearing) transformation and translation, $\mathcal{T}|_{\theta_\phi}$ refers to a non-linear displacement field computed as the integration of a stationary velocity field (SVF) that is smooth and invertible everywhere and thus preserves the topology of the brain anatomy [28]. $\theta_\phi \in \Theta$ controls the transformation ranges.

MR Contrast Synthesis Then, we synthesize images ($I(x)$, $x \in \Omega$) by randomly "painting" intensities on the segmentation maps according to their brain structure labels ($l \in L$). Specifically, the regional intensities are generated by separately sampling a Gaussian distribution on each labeled region:

$$\begin{cases} I(x) \sim \mathcal{N}(\mu_l,\, \sigma_l)\,, \quad l \in L\,, \\ \mu_l \sim \mathcal{N}(0,\, 1\,|\,\theta_\mu,\, \theta_l)\,,\ \sigma_l \sim \mathcal{N}(0,\, 1\,|\,\theta_\sigma,\, \theta_l)\,, \end{cases} \tag{2}$$

where μ_l and σ_l refer to the mean and standard deviation of each segmentation label l, and are independently sampled from Gaussian distributions at each voxel, with θ_l, θ_μ, $\theta_\sigma \in \Theta$ controlling the shifts and scales of their values.

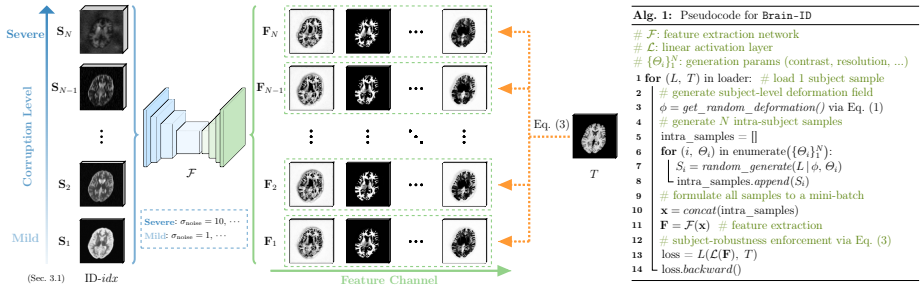

Fig. 3: Brain-ID's contrast-agnostic anatomical representation learning framework.

Resolution Simulation and Data Corruption Given a deformed, contrast-synthesized image (I), we adopt the standard data corruption pipeline [27], which further augments images with different levels of resolution and scanning artifacts that are commonly found in real-world clinical protocols.

As illustrated in Fig. 2, we are able to generate an infinite number of variations from a single subject with its unique brain anatomy. Generating images with randomized contrast/resolution/orientation on the fly for each subject enormously enriches the learning space for a robust representation, and helps focus the learning on intrinsic subject-specific features rather than superficial aspects of images that depend on acquisition parameters and conditions.

3.2 Extract Robust and Expressive Subject-specific Features

As mentioned above, we would like Brain-ID features to be robust to intra-subject variations *and* expressive to potential downstream tasks. In this section, we introduce Brain-ID's learning framework to achieve the two goals.

Intra-subject Data Generation In order to learn a feature representation that is distinctive to each subject and robust to varying MR contrasts, we observe that enriching *intra*-subject samples leads to better performance (Tab. 3). Specifically, instead of including multiple subjects for a mini-batch during each training iteration as in usual practice, Brain-ID focuses on maximizing the intra-subject variance to improve the subject-robustness of resulting features. As described in Alg. 1 (line 4-10), for each training iteration, after randomly selecting a subject and generating its deformation, Brain-ID generates a mini-batch of intra-subject samples ($\{S_1, \ldots, S_N\}$) with randomly synthesized contrasts, resolutions and corruptions (Sec. 3.1). As will be introduced below, Brain-ID collects losses from all intra-subject samples and conducts back-propagation *at once*, to encourage the subject-specific robustness of its learned features.

We set the intra-subject samples within a mini-batch to have random contrasts and "mild-to-severe", *increasing* level of corruptions (Fig. 3 (left)), to maximize the intra-subject variance while ensuring the stability of the training process against extreme corruption levels. (In Sec. 4.5, we compare various data generation designs, and provide insights on preventing unstable training.)

Anatomy-guided Feature Representation A richer learning space helps improve the representation robustness against the variance of sample appearances, but not enough for extracting expressive features. For example, a mapping that projects all inputs to zero is perfectly robust, but does not provide useful information for potential downstream tasks. Therefore, proper guidance is crucial for feature representation as well. In CIFL [15], the authors use a contrastive loss on feature channels to encourage a more discriminative feature representation. However, this method is insufficient, given the extremely complex nature of the human brain anatomy (See comparisons in Sec. 4). Instead, we propose to use the standard T1w high-resolution structural MR contrast for brain morphometry, i.e., MP-RAGE (magnetization-prepared rapid gradient-echo), as the unique anatomy target to guide `Brain-ID`'s feature representation learning.

As shown in Fig. 3, a feature extraction backbone ($\mathcal{F}$) first maps the input mini-batch of intra-subject generated samples, $\{S_1, \ldots, S_N\}$, to their corresponding feature space, $\{\mathbf{F}_1, \ldots, \mathbf{F}_N\}$. A linear activation layer ($\mathcal{L}$) then projects the features to the current subject's standard contrast (MP-RAGE) space, T. The training loss is obtained by summing over the anatomy reconstruction loss of all intra-subject samples in the current mini-batch:

$$L = \sum_i^N \; |\mathcal{L}(\mathbf{F}_i) - T| + \lambda \, |\nabla\mathcal{L}(\mathbf{F}_i) - \nabla T| \,, \; \lambda \in R^+ \,, \tag{3}$$

where T, as the high-resolution, unique anatomy target for all intra-subject samples, encourages the *similarities* of the features extracted from the same subject via reconstruction (1^{st} term) and gradient difference (2^{nd} term) `L1` loss. `Brain-ID` uses MP-RAGE as the learning target, which formulates both super-resolution and contrast-synthesis problems, and encourages a richer feature representation. (Sec. 4.5 provides further insights on how different choices of anatomy guidance would affect the resulting features' properties.)

3.3 Adapting `Brain-ID` to Downstream Tasks by One Layer

With the intra-subject data generation and anatomy-guided feature learning design introduced in Sec. 3.2, a well-trained `Brain-ID` model is able to extract robust, high-resolution anatomical features from images with varying deformations, contrasts, resolutions, and artifacts. To minimize the modifications, we propose straightforward *one-layer* solutions adapting `Brain-ID` features to various brain imaging applications. We later demonstrate that the simple adaptions are effective enough for `Brain-ID` to achieve state-of-the-art performance across all downstream tasks (Sec. 4.3), even for small datasets (Sec. 4.4).

Contrast-independent Tasks For tasks where the output should be independent of the input MR contrast, e.g., brain segmentation, we simply add an additional layer following `Brain-ID` features ($\mathbf{F}$), and fine-tune the model:

$$L = task\_loss\_func\big(\mathcal{L}(\mathbf{F}), T\big) \,, \tag{4}$$

where $\mathcal{L}$ and T refer to the task-specific activation layer and the contrast-independent ground truth, respectively.

Contrast-dependent Tasks Since Brain-ID features are contrast-agnostic and robust to artifacts, for tasks relevant to the input's contrast/quality, e.g., super-resolution, we concatenate the input image (I) that contains the original contrast information with its high-resolution Brain-ID features along the channel dimension, before forwarding into the task-specific activation layer:

$$L = task\_loss\_func\big(\mathcal{L}(\mathbf{F} \oplus I),\, T\big)\,, \tag{5}$$

where $\mathcal{L}$ and T refer to the task-specific activation layer and the ground truth.

4 Experiments

We demonstrate the two properties of Brain-ID as claimed in Sec. 3. *(i)* Robustness (Sec. 4.2): we propose "canonical" and "atlas-registered" features to assess the intra/inter-subject feature robustness; *(ii)* Expressiveness (Sec. 4.3): we evaluate the performance of adapting Brain-ID features to various brain imaging applications. We further challenge Brain-ID with reduced-size datasets to explore its ability when limited real data is available for training (Sec. 4.4).

4.1 Datasets, Metrics and Implementation Details

Datasets For Brain-ID features pre-training, we use 2045 3D segmentations ("anatomies", or IDs) from the public ADNI dataset [30]. For downstream evaluation, we use six public datasets covering T1w and T2w MRI, FLAIR MRI, and CT: ADNI [30], ADNI3 [60], HCP [19], ADHD200 [9], AIBL [23], OASIS3 [35].

Metrics We evaluate individual tasks from different aspects. For feature similarity measurements, we use L1 distance, and (MS-)SSIM (multi-scale structural similarity) [41, 58, 59]. For reconstruction and super-resolution, we use L1, PSNR (peak signal-to-noise ratio) and (MS-)SSIM. For segmentation, we use Dice scores. For bias field estimation, we use the normalized L2 distance (norm-L2) to avoid possible arbitrary scalings from nonuniformity correction [16].

Implementation Details As a general feature representation model, Brain-ID can use any backbone to extract brain features. For fairer comparison, we adopt the same five-level 3D UNet [47] (with 64 feature channels in the last layer) as utilized in state-of-the-art models we compare with in Sec. 4.3. During feature pre-training (Sec. 3.2), a linear regression layer is added for anatomy supervision. For downstream adaptions, A task-specific adaption layer is added (Sec. 3.3).

4.2 Intra/Inter-subject Feature Robustness

In this section, we examine the robustness of Brain-ID features, i.e., the 64 features from the last layer. For comparable and reproducible results, we use our data generator (Sec. 3.1) to prepare (deform and corrupt) 1000 testing samples from 100 randomly selected subjects in ADNI [30] (T1w) testing set, with 10 intra-subject samples for each subject.

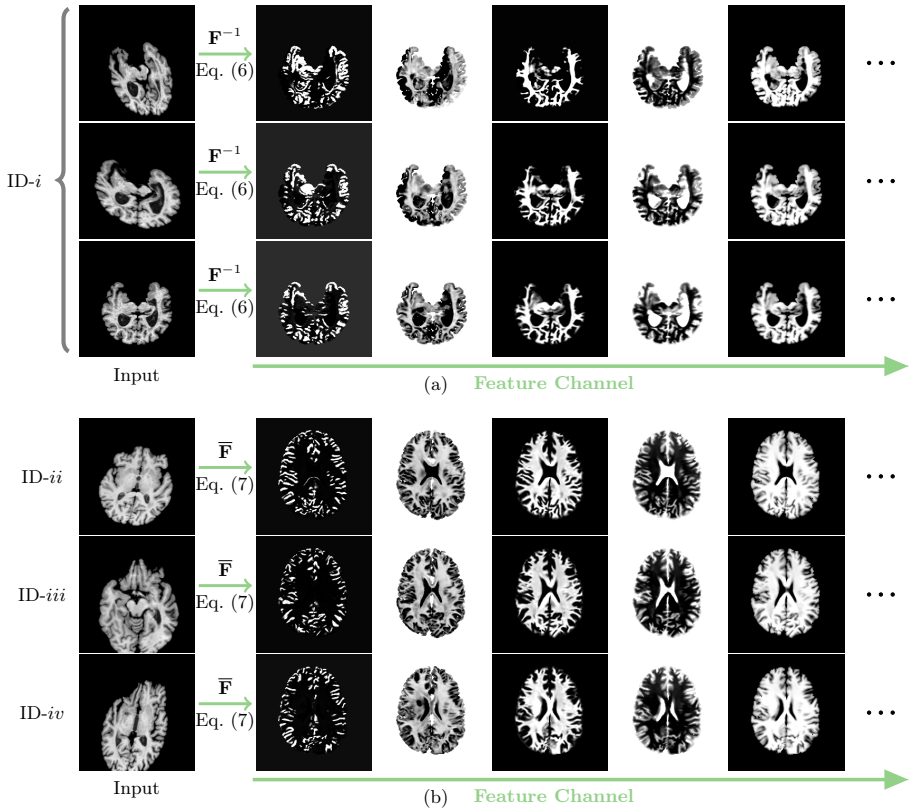

Fig. 4: (a) Intra-subject and (b) inter-subject robustness of `Brain-ID` features.

We compare our features with `CIFL` [15] which is, to our best knowledge, the only similar work to `Brain-ID`, that also aims to learn contrast-agnostic brain features. Note the original `CIFL` method has only experimented on 2D images, and does not have our intra-subject data generation design. For a fairer comparison, we trained `CIFL` *with its contrastive-learning design*, while using the same data generator and network architecture as `Brain-ID`.

Intra-subject Feature Robustness Ideally, `Brain-ID` features computed from the same subject should be structurally *identical*, regardless of their appearances (contrast, resolution, noises, etc). To assess the intra-subject feature robustness, we first define their corresponding "canonical" features.

"Canonical" Features. We compute ϕ^{-1}, the inverse of S' generated deformation as in Eq. (1), and map back each sample's (S) `Brain-ID` features ($\mathbf{F} = \mathbf{F}(S \,|\, \phi)$) to their original domain without any deformation:

$$\mathbf{F}^{-1} = \phi^{-1} \circ \mathbf{F}(S \,|\, \phi)\,. \tag{6}$$

Table 1: Evaluations on intra/inter-subject feature robustness.

Mode	Method	L1 (↓)	SSIM (↑)	MS-SSIM (↑)
Intra	CIFL [15]	0.122 (±0.031)	0.511 (±0.298)	0.531 (±0.367)
	Brain-ID	**0.011** (±0.001)	**0.858** (±0.041)	**0.921** (±0.008)
Inter	CIFL [15]	0.230 (±0.039)	0.552 (±0.240)	0.524 (±0.291)
	Brain-ID	**0.014** (±0.013)	**0.843** (±0.044)	**0.913** (±0.011)

As shown in Fig. 4-(a), for the same subject (ID-i), although the inputs (row) are processed by varying deformations and corruptions, their Brain-ID's "canonical" features are of similar looking along each feature channel (column). In Tab. 1, we quantify the feature similarity: For each testing subject, we first generate a sample *without* any deformation or corruption and use its Brain-ID feature as the gold-standard reference for all other "canonical" features computed from intra-subject samples that may have different levels of deformation and corruptions. The final results are obtained by scores averaged over all 64 feature channels of all the intra-subject samples.

Inter-subject Feature Robustness We further assess the feature robustness among anatomies, based on the assumption that features from different subjects should ideally be *structurally similar*, once being registered to a common domain. *"Atlas-registered" Features.* We register [43] features ($\mathbf{F} = \mathbf{F}(S \,|\, \phi)$) from all subjects to a common brain atlas [21, 22]:

$$\overline{\mathbf{F}} = \psi_{S \to A} \circ \mathbf{F}(S \,|\, \phi)\,, \tag{7}$$

where $\psi_{S \to A}$ denotes the mapping from S to the atlas, A.

As shown in Fig. 4-(b), even though the input samples (row) from different subjects have varying anatomical structures, after being registered to a common domain, their "atlas-registered" Brain-ID features still appear consistent (column). In Tab. 1, we further provide quantitative results of the structural similarity measured between features from different subjects. Similar to Sec. 4.2, we randomly select a testing subject and generate a sample *without* any deformation or corruption. We use its "atlas-registered" Brain-ID feature as the gold standard for all "atlas-registered" features computed from other samples that may have different levels of deformation and corruption. The final results are obtained by scores averaged over all samples of the 100 testing subjects.

In both intra/inter-subject aspects, Brain-ID outperforms CIFL by a large margin, which validates its feature robustness against deformations, contrasts, resolutions, artifacts (intra-subject), and anatomies (inter-subject).

4.3 Downstream Evaluation

Following the standard evaluation protocol for representation learning [12], we fine-tune Brain-ID features for downstream tasks through one-layer (Sec. 3.3).

Models We compare the downstream performance of Brain-ID features with: *(i-ii)* SAMSEG [13, 46] and FastSurfer [24] (only works on T1w), state-of-the-art

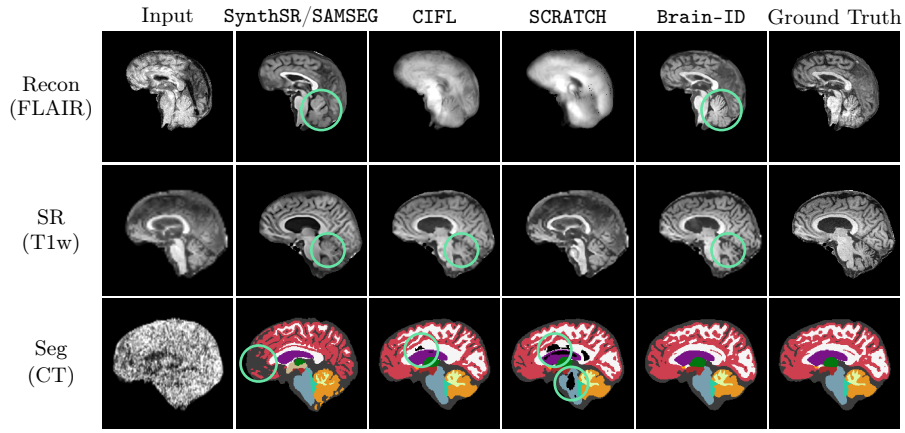

Fig. 5: Qualitative comparisons on downstream tasks of reconstruction (Recon), super-resolution (SR), and segmentation (Seg), between `Brain-ID`, the baseline `SCRATCH`, and the state-of-the-art methods `CIFL` [15], `SynthSR` [27] (Recon and SR), `SAMSEG` [13] (Seg). The visualized testing examples are taken from: AIBL-FLAIR [23] for Recon, AIBL-T1w [23] for SR, and OASIS-CT [35] for Seg. The mint circles highlight some details.

classical and machine-learning-based brain segmentation models, respectively; *(iii)* `SynthSR` [27], state-of-the-art, contrast-agnostic T1w synthesis model. We also provide fine-tuned `SynthSR` (`SynthSR`-FT) results for further comparison; *(iv)* `SCRATCH`: a baseline with the same data generation/architecture as `Brain-ID`, yet trained from *scratch*. `SCRATCH` is used to validate `Brain-ID`'s effectiveness. *(v)* `CIFL` [15]: a baseline with the same data generation/architecture as `Brain-ID` and `SCRATCH`, yet initialized with `CIFL`'s pre-trained features. We use `CIFL` to demonstrate the superiority of `Brain-ID`'s anatomy-guided representation.

Contrast-independent Applications

- Anatomy Reconstruction/Contrast Synthesis As one of the most common types of MRI scans, T1w MR images highlight the differences between gray and white matter, and are mostly often utilized to image the anatomy of the brain, spinal cord, bones and joints [51]. For each dataset, we train models to reconstruct their T1w MRI counterparts with the `L1` loss.

- Brain Segmentation For each dataset, we use SynthSeg [4] to obtain the segmentation labels with 30 brain anatomical regions [20], as the gold standard segmentation target. We train our compared models to predict the brain segmentation labels, with the soft dice loss and cross-entropy loss [4].

As shown in Fig. 5, `Brain-ID` clearly outperforms `CIFL` and `SCRATCH` in the reconstruction task. As highlighted by the mint circles, it reveals the anatomy details and produces more fine-grained results than the strong state-of-the-art method `SynthSR` [27]. `Brain-ID` also achieves better segmentation results, especially in the smaller and more challenging regions. Quantitative comparisons are in Tab. 2, where `Brain-ID` obtains the best scores across most metrics. No-

Table 2: Comparisons of Brain-ID with state-of-the-art approaches on downstream applications. ("train/test" refers to the number of subjects in the training/testing set.)

Modality (Contrast)	Dataset (Train/Test)	Method	Reconstruction				Segmentation	Super-resolution (from 3 / 5 / 7 *mm*)			Bias Field
			L1 (↓)	PSNR (↑)	SSIM (↑)	MS-SSIM (↑)	Dice (↑)	PSNR (↑)	SSIM (↑)	MS-SSIM (↑)	norm-L2 (↓)
MR (T1w)	*ADNI* [30] (1841/204)	SAMSEG [13]	-	-	-	-	0.795	-	-	-	-
		FastSurfer [24]	-	-	-	-	0.803	-	-	-	-
		SynthSR [27]	0.014	26.78	0.980	0.980	-	27.01 / 25.81 / 21.79	0.977 / 0.943 / 0.857	0.981 / 0.968 / 0.869	-
		SynthSR-FT	**0.010**	29.54	0.984	0.920	0.825	30.46 / 28.03 / 23.91	0.979 / 0.960 / 0.898	0.985 / 0.973 / 0.900	0.051
		CIFL [15]	0.013	26.97	0.973	0.961	0.820	30.12 / 27.21 / 23.80	0.978 / 0.912 / 0.887	0.979 / 0.950 / 0.842	0.053
		SCRATCH	0.011	27.54	0.979	0.957	0.816	30.29 / 27.94 / 24.57	0.975 / 0.959 / 0.920	0.982 / 0.964 / 0.866	0.056
		Brain-ID	0.012	**33.82**	**0.993**	**0.989**	**0.837**	**31.30** / **28.62** / **24.68**	**0.983** / **0.961** / **0.928**	**0.987** / **0.980** / **0.947**	**0.048**
	HCP [19] (808/87)	SAMSEG [13]	-	-	-	-	0.810	-	-	-	-
		FastSurfer [24]	-	-	-	-	0.819	-	-	-	-
		SynthSR [27]	0.033	22.13	0.854	0.901	-	22.21 / 20.83 / 19.93	0.848 / 0.802 / 0.747	0.899 / 0.864 / 0.789	-
		SynthSR-FT	0.025	27.11	0.935	0.917	0.810	25.90 / 25.45 / 25.10	0.889 / 0.872 / 0.831	0.919 / 0.911 / 0.889	0.052
		CIFL [15]	0.029	26.42	0.932	0.913	0.804	25.98 / 25.11 / 23.99	0.930 / 0.883 / 0.871	0.936 / 0.927 / 0.880	0.059
		SCRATCH	0.023	27.32	0.923	0.909	0.792	26.69 / 26.41 / 25.05	0.932 / 0.896 / 0.848	0.942 / 0.930 / **0.892**	0.069
		Brain-ID	**0.020**	**27.47**	**0.957**	**0.929**	**0.843**	**29.70** / **27.90** / **26.12**	**0.957** / **0.901** / **0.883**	**0.973** / **0.938** / 0.890	**0.047**
	ADNI3 [60] (298/33)	SAMSEG [13]	-	-	-	-	0.769	-	-	-	-
		FastSurfer [24]	-	-	-	-	0.796	-	-	-	-
		SynthSR [27]	0.023	23.60	0.928	0.909	-	23.13 / 22.40 / 22.27	0.921 / 0.907 / 0.876	0.903 / 0.892 / 0.871	-
		SynthSR-FT	0.021	27.11	0.950	0.969	0.820	27.95 / 27.34 / 26.18	0.920 / 0.913 / 0.890	0.922 / 0.890 / 0.879	0.134
		CIFL [15]	0.028	28.52	0.961	0.970	0.819	27.81 / 27.10 / 26.32	0.900 / 0.877 / 0.864	0.812 / 0.793 / 0.769	0.102
		SCRATCH	0.033	27.28	0.957	0.963	0.816	27.99 / 27.37 / 26.63	0.893 / 0.889 / 0.876	0.794 / 0.785 / 0.753	0.128
		Brain-ID	**0.021**	**29.89**	**0.966**	**0.976**	**0.843**	**30.17** / **28.23** / **26.89**	**0.962** / **0.923** / **0.892**	**0.973** / **0.921** / **0.883**	**0.093**
	ADHD200 [9] (865/96)	SAMSEG [13]	-	-	-	-	0.784	-	-	-	-
		FastSurfer [24]	-	-	-	-	0.801	-	-	-	-
		SynthSR [27]	0.035	21.67	0.882	0.853	-	21.42 / 21.13 / 21.41	0.876 / 0.846 / 0.809	0.851 / 0.831 / 0.805	-
		SynthSR-FT	0.016	26.90	0.917	0.905	0.812	28.55 / 27.43 / 25.94	0.880 / 0.859 / 0.823	0.860 / 0.845 / 0.811	0.110
		CIFL [15]	0.013	28.69	0.932	0.921	0.810	28.11 / 26.12 / 25.11	0.865 / 0.848 / **0.831**	0.828 / 0.800 / 0.789	0.109
		SCRATCH	**0.011**	31.55	0.986	0.985	0.796	29.41 / 27.89 / 26.58	0.858 / 0.847 / 0.814	0.853 / 0.845 / **0.827**	0.113
		Brain-ID	**0.011**	**32.48**	**0.996**	**0.996**	**0.847**	**29.50** / **28.56** / **26.87**	**0.898** / **0.862** / 0.811	**0.887** / **0.850** / 0.823	**0.107**
	AIBL [23] (601/67)	SAMSEG [13]	-	-	-	-	0.799	-	-	-	-
		FastSurfer [24]	-	-	-	-	0.802	-	-	-	-
		SynthSR [27]	0.026	22.95	0.916	0.912	-	22.57 / 21.82 / 21.78	0.917 / 0.892 / 0.856	0.906 / 0.893 / 0.869	-
		SynthSR-FT	0.014	29.89	0.941	0.922	0.810	29.15 / 27.89 / 26.77	0.932 / 0.906 / 0.878	0.927 / 0.900 / 0.871	0.118
		CIFL [15]	0.011	30.12	0.938	0.925	0.816	29.46 / 27.97 / 26.73	0.907 / 0.889 / 0.870	0.891 / 0.872 / 0.857	0.094
		SCRATCH	0.011	30.53	0.932	0.913	0.814	29.43 / 28.10 / 26.82	0.900 / 0.886 / 0.863	0.884 / 0.870 / 0.853	0.128
		Brain-ID	**0.009**	**33.73**	**0.972**	**0.963**	**0.851**	**29.75** / **28.58** / **27.09**	**0.957** / **0.922** / **0.890**	**0.975** / **0.956** / **0.937**	**0.088**
MR (T2w)	*HCP* [19] (808/87)	SAMSEG [13]	-	-	-	-	0.782	-	-	-	-
		SynthSR [27]	0.034	21.46	0.833	0.885	-	-	-	-	-
		SynthSR-FT	0.028	26.10	0.894	0.898	0.755	28.43 / 25.36 / 22.10	0.942 / 0.851 / 0.801	0.977 / 0.892 / 0.798	0.140
		CIFL [15]	0.036	25.12	0.891	0.879	0.787	28.26 / 25.78 / 21.73	**0.959** / 0.893 / 0.797	0.982 / 0.906 / 0.742	0.138
		SCRATCH	0.038	24.99	0.873	0.866	0.756	28.12 / 24.52 / 21.58	0.945 / 0.863 / 0.783	0.980 / 0.893 / 0.738	**0.136**
		Brain-ID	**0.016**	**28.10**	**0.934**	**0.935**	**0.805**	**30.26** / **26.11** / **24.10**	**0.959** / **0.902** / **0.832**	**0.987** / **0.953** / **0.912**	**0.136**
	AIBL [23] (272/30)	SAMSEG [13]	-	-	-	-	0.763	-	-	-	-
		SynthSR [27]	0.033	20.08	0.805	0.831	-	-	-	-	-
		SynthSR-FT	0.024	22.93	0.815	0.840	0.719	30.15 / 28.09 / 26.92	0.940 / 0.912 / 0.893	0.966 / 0.948 / 0.922	0.200
		CIFL [15]	0.022	23.71	0.820	0.839	0.721	29.98 / 27.01 / 25.66	0.931 / 0.899 / 0.861	0.967 / 0.941 / 0.905	0.193
		SCRATCH	**0.020**	22.27	0.849	0.851	0.714	31.91 / 29.20 / **27.35**	0.954 / 0.934 / 0.896	0.982 / 0.969 / 0.947	0.197
		Brain-ID	0.022	**23.99**	**0.861**	**0.850**	**0.782**	**32.26** / **29.90** / 27.09	**0.973** / **0.937** / **0.900**	**0.988** / **0.971** / **0.948**	**0.148**
MR (FLAIR)	*ADNI3* [60] (298/33)	SAMSEG [13]	-	-	-	-	0.718	-	-	-	-
		SynthSR [27]	0.026	22.77	0.919	0.895	-	-	-	-	-
		SynthSR-FT	0.021	22.34	0.921	**0.900**	0.753	30.11 / 28.70 / 25.32	0.930 / 0.899 / 0.864	0.948 / 0.910 / 0.878	0.251
		CIFL [15]	0.020	21.29	0.911	0.897	0.761	**32.72** / 29.00 / 26.99	0.949 / 0.906 / 0.873	0.953 / 0.919 / 0.889	0.237
		SCRATCH	0.025	20.80	0.901	0.862	0.759	32.36 / 28.71 / 28.00	0.945 / 0.915 / 0.877	0.941 / 0.936 / 0.917	0.236
		Brain-ID	**0.017**	**26.44**	**0.927**	0.892	**0.786**	32.18 / **30.00** / **28.28**	**0.959** / **0.921** / **0.883**	**0.982** / **0.965** / **0.921**	**0.227**
	AIBL [23] (302/34)	SAMSEG [13]	-	-	-	-	0.710	-	-	-	-
		SynthSR [27]	0.029	21.77	0.902	0.892	-	-	-	-	-
		SynthSR-FT	0.024	25.80	0.914	0.900	0.735	26.20 / 24.91 / 23.17	0.860 / 0.832 / 0.795	0.933 / 0.910 / 0.894	0.126
		CIFL [15]	0.026	27.11	0.901	0.870	0.721	27.72 / 26.85 / 25.99	0.851 / 0.828 / 0.804	0.939 / 0.905 / 0.901	0.109
		SCRATCH	0.023	26.84	0.898	0.867	0.720	28.02 / 26.52 / 25.43	0.878 / 0.838 / 0.799	0.943 / 0.901 / 0.891	0.115
		Brain-ID	**0.019**	**27.25**	**0.936**	**0.912**	**0.767**	**28.69** / **27.63** / **26.69**	**0.949** / **0.906** / **0.866**	**0.971** / **0.925** / **0.916**	**0.103**
CT	*OASIS3* [35] (795/88)	SAMSEG [13]	-	-	-	-	0.701	-	-	-	-
		SynthSR [27]	0.041	20.93	0.758	0.786	-	-	-	-	-
		SynthSR-FT	0.030	23.76	0.797	0.845	0.700	25.30 / 22.18 / 20.35	0.892 / 0.813 / 0.760	0.896 / 0.801 / 0.764	-
		CIFL [15]	0.027	24.91	0.819	0.871	0.718	25.99 / 23.70 / 22.83	0.909 / **0.820** / 0.779	0.969 / 0.816 / 0.775	-
		SCRATCH	0.025	24.35	0.811	0.872	0.709	26.34 / 23.64 / 22.62	0.905 / 0.818 / 0.770	0.873 / 0.818 / 0.780	-
		Brain-ID	**0.023**	**25.49**	**0.891**	**0.895**	**0.765**	**26.74** / **24.01** / **23.09**	**0.910** / 0.818 / **0.792**	**0.974** / **0.824** / **0.799**	-

(1) "-" means not applicable; FastSurfer [24] and SAMSEG [13] are for segmentation-only; SynthSR [27] only synthesize T1w MRI; CT does not have bias fields.
(2) SynthSR-FT: SynthSR fine-tuned on each task/dataset, for a fairer comparison.

tably, Brain-ID achieves greater gains over baseline models, SCRATCH and CIFL, particularly on smaller datasets (e.g., ADNI3-T1/FLAIR, AIBL-FLAIR, where Brain-ID's robust and rich features quickly adapt to specific tasks/datasets.

Contrast-dependent Applications We also evaluate Brain-ID on two brain imaging tasks that are dependent on the input modality/quality.

- Image Super-resolution We use the standard $1mm$ isotropic images from all datasets as the ground truth high-resolution target, and follow the strategy in SynthSR [27] where the input samples are randomly resampled and corrupted during training. For inference, we downsample the original images into $3mm$, $5mm$, and $7mm$ isotropic images as inputs, and evaluate the output quality.

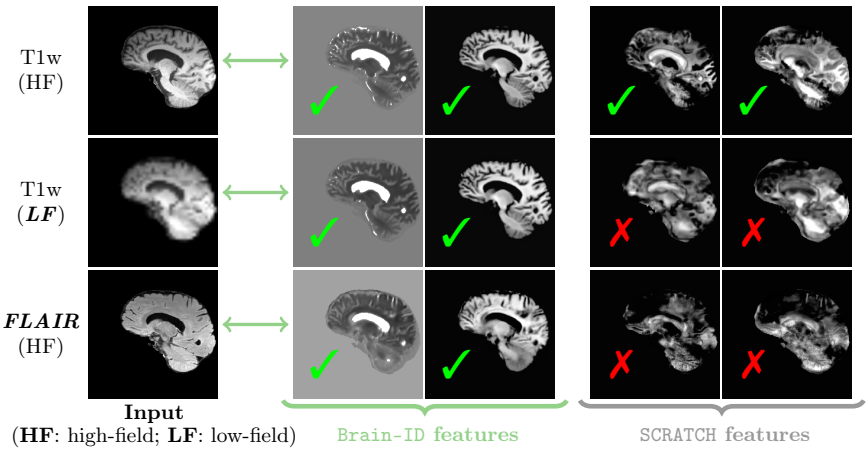

Fig. 6: SCRATCH, which is well trained on HF T1w scans, produces highly descriptive features for HF T1w images (1st row), but *does not preserve* the same high quality useful for downstream tasks when handling LF (2nd row) or other contrasts (3rd row).

- Bias Field Estimation The bias field is a smooth, low-frequency multiplicative signal that corrupts MRI images, which affects image analysis tasks such as segmentation or texture analysis [31]. Bias field estimation is often needed as a pre-processing step to correct corrupted MRI images [16]. We apply randomly simulated bias fields to the input samples (Sec. 3.1), and train all models with the L2 loss. During inference, we pre-generate and apply the bias fields on the testing data for reproducibility, and evaluate the bias field estimation performance.

With a simple one-layer adaption (Sec. 3.3), Brain-ID's contrast-agnostic anatomical features achieve state-of-the-art performance on contrast-dependent tasks (Tab. 2, Fig. 5). As shown in Fig. 6, Brain-ID obtains contrast/resolution-robustness that *cannot* be achieved by models trained from real images (due to the *limited variability* in their appearance), regardless of the backbone choices. Such robust features greatly improve super-resolution, resulting in higher (MS-)SSIM scores. Brain-ID's gains are less obvious for bias field estimation, probably because the bias field is approximately *independent* of brain anatomy.

4.4 Practical Value and Broader Impact

Low-resolution data Trained on synthetic images, Brain-ID features are not only robust to various contrasts/modalities, but also extremely robust to low-resolution data (Fig. 6) and provide huge potential in clinical MRI (including big retrospective data [5]) and portable low-field MRI [29].

Small-size Datasets To investigate the effectiveness of Brain-ID features for limited-size datasets, we assess its performance across all four downstream tasks on subsets of ADNI3 [60] training set (298 cases originally). As shown in Fig. 7, we reduce the training set size percentage from 100% to an extreme 10%. With

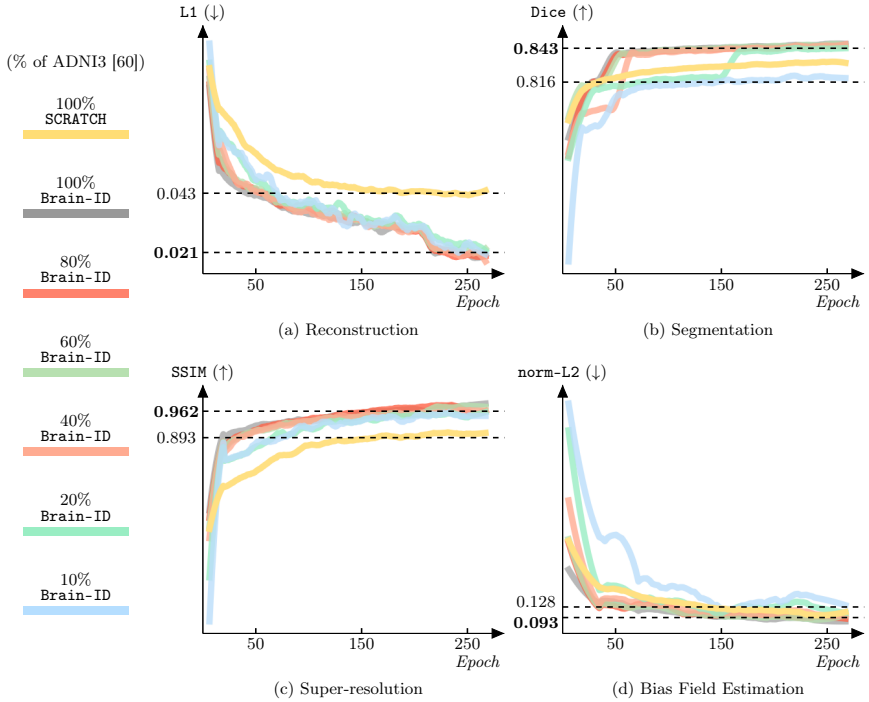

Fig. 7: Adapting Brain-ID to small datasets. The horizontal (vertical) axes indicate training epochs (evaluation scores of corresponding tasks). Results are obtained by evaluating models collected throughout the epochs, on ADNI3 [60] full testing set.

only 20% of the data (mint line), Brain-ID still achieves better performance compared to its full-sized baseline, "100%" SCRATCH (yellow line), over all tasks. 30 training samples (10%, blue line) may be at the edge of obtaining acceptable results for Brain-ID, as the model becomes less stable and the performance drops occur more frequently – especially for brain segmentation (Fig. 7-(b)).

4.5 Additional Experiments

Choice of Anatomy Guidance We explore other anatomical learning targets than MP-RAGE (Sec. 3.2): *(i)* segmentation labels; *(ii)* both segmentation labels and MP-RAGE. As shown in Tab. 3, adding segmentation guidance improves robustness, yet inhibits the feature expressiveness and affects the downstream performance; we observe features from segmentation guidance produce less high-frequency texture than Brain-ID, especially *within* the same-label regions.

Data Generation Design As shown in Tab. 3, training with all mildly-corrupted samples results in reduced robustness and harms the downstream performance, yet using all extremely corrupted samples leads to unstable training and model collapse. Brain-ID uses samples of gradually increased corruptions (Fig. 3 (left)).

Table 3: Comparison between `Brain-ID` and its variants.

Model Setup	Feature Robustness (Intra)		Downstream (Reconstruction)		
[Comparison Target]	`SSIM` (↑)	`MS-SSIM` (↑)	`L1` (↓)	`PSNR` (↑)	`SSIM` (↑)
[Representation guidance: supervision target for feature learning]					
Segmentation	**0.891**	**0.963**	0.029	28.13	0.958
Segmentation + MP-RAGE	0.863	0.940	0.023	29.05	0.964
MP-RAGE (`Brain-ID`)	0.858	0.921	**0.021**	**29.89**	**0.966**
[Data generation design: corruption levels in intra-subject mini-batches]					
All Mild ($\sigma_{\text{noise}} = 1, \cdots$)	0.792	0.813	0.028	28.30	0.960
All Medium ($\sigma_{\text{noise}} = 5, \cdots$)	0.831	0.899	0.022	29.67	0.964
All Severe ($\sigma_{\text{noise}} = 10, \cdots$)	N/A	N/A	N/A	N/A	N/A
Mild to Severe (`Brain-ID`)	**0.858**	**0.921**	**0.021**	**29.89**	**0.966**
[Mini-batch size: number of intra-subject samples]					
2	0.826	0.897	0.025	28.83	0.959
3	0.841	0.902	0.022	29.30	0.962
4 (`Brain-ID`)	0.858	0.921	**0.021**	29.89	**0.966**
5	**0.860**	**0.929**	**0.021**	**29.93**	0.965

Batch Size `Brain-ID` computes training loss of all intra-subject samples in a mini-batch at once (Eq. (3)). We observe that larger batches help improve feature robustness (Tab. 3), yet do not further enhance the downstream performance.

4.6 Further Discussions

Limitations Although `Brain-ID` features excel in multiple tasks, they may not be optimal for tasks not involving voxel-wise prediction, e.g., functional brain mapping. Further, as `Brain-ID` generates data exclusively from anatomies, it tends to struggle with images with extensive pathology. Our future research will focus on robust representations that are encoded with pathology information.

Data Privacy and Ethics Since `Brain-ID` features can potentially be used to reconstruct a high-resolution MRI scan, publicly sharing such features should be regarded as sharing the original scans from a subject [33]. Therefore, one should follow the same anonymization protocols as when sharing MRI data for research [61] – not only the removal of identifiers, but also defacing [6] or skull-stripping [26] to prevent the 3D reconstruction of faces.

5 Conclusion

We introduced `Brain-ID`, a contrast-agnostic feature representation learning approach for brain imaging, which is distinctive to brain anatomy and resilient to variations in image appearances. `Brain-ID` is trained on synthetic inputs, and quickly adapts to downstream tasks through one layer. We validated `Brain-ID` features' robustness, and effectiveness on four downstream applications (reconstruction, segmentation, super-resolution, bias field estimation), and six public datasets, covering MR and CT. `Brain-ID` achieves state-of-the-art performances across all tasks and modalities, it further preserves high performance on low-resolution and small-size datasets. We believe `Brain-ID` unlocks the great potential of robust foundation models, especially for non-calibrated modalities.

Acknowledgements

Primarily supported by NIH under grant 1RF1AG080371. Additional support from NIH 1RF1AG080371, 1R01EB031114, 1R01AG070988, 1RF1MH123195. OP is supported by Lundbeckfonden R360–2021–395. Wellcome Trust award 221915/Z/20/Z and the NIHR UCLH Biomedical Research Centre support DCA on this topic.

References

1. Adnan, M., Kalra, S., Cresswell, J.C., Taylor, G.W., Tizhoosh, H.R.: Federated learning and differential privacy for medical image analysis. Scientific reports (2022)
2. Awais, M., Naseer, M., Khan, S.S., Anwer, R.M., Cholakkal, H., Shah, M., et al.: Foundational models defining a new era in vision: A survey and outlook. arXiv **abs/2307.13721** (2023)
3. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J.V., Dalca, A.V.: VoxelMorph: A learning framework for deformable medical image registration. IEEE Transactions on Medical Imaging (2018)
4. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Leemput, K.V., Fischl, B., et al.: SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. Medical Image Analysis (2021)
5. Billot, B., Magdamo, C., Cheng, Y., Arnold, S.E., Das, S., Iglesias, J.E.: Robust machine learning segmentation for large-scale analysis of heterogeneous clinical brain MRI datasets. Proceedings of the National Academy of Sciences (2023)
6. Bischoff-Grethe, A., Ozyurt, I.B., Busa, E., Quinn, B.T., Fennema-Notestine, C., Clark, C.P., et al.: A technique for the deidentification of structural brain MR images. Human Brain Mapping (2007)
7. Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al.: On the opportunities and risks of foundation models. arXiv **abs/2108.07258** (2021)
8. Brant-Zawadzki, M.N., Gillan, G.D., Nitz, W.R.: MP RAGE: A three-dimensional, T1-weighted, gradient-echo sequence–initial experience in the brain. Radiology (1992)
9. Brown, M.R.G., Sidhu, G.P.S., Greiner, R., Asgarian, N., Bastani, M., Silverstone, P.H., et al.: ADHD-200 global competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements. Frontiers in Systems Neuroscience (2012)
10. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., et al.: Language models are few-shot learners. In: NeurIPS (2020)
11. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., et al.: MONAI: An open-source framework for deep learning in healthcare. arXiv **abs/2211.02701** (2022)
12. Caron, M., Touvron, H., Misra, I., J'egou, H., Mairal, J., Bojanowski, P., Joulin, A.: Emerging properties in self-supervised vision transformers. In: ICCV (2021)
13. Cerri, S., Puonti, O., Meier, D.S., Wuerfel, J., Mühlau, M., Siebner, H.R., Leemput, K.V.: A contrast-adaptive method for simultaneous whole-brain and lesion segmentation in multiple sclerosis. NeuroImage (2020)

14. Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., et al.: PaLM: Scaling language modeling with pathways. JMLR (2022)
15. Chua, Y.Z.R., Dalca, A.V.: Contrast invariant feature representations for segmentation and registration of medical images. In: MIDL (2023)
16. Chua, Z.Y., Zheng, W., Chee, M.W.L., Zagorodnov, V.: Evaluation of performance metrics for bias field correction in MR brain images. Journal of Magnetic Resonance Imaging (2009)
17. Darzidehkalani, E., Ghasemi-Rad, M., van Ooijen, P.: Federated learning in medical imaging: Part I: Toward multicentral health care ecosystems. Journal of the American College of Radiology (2022)
18. Ding, Z., Han, X., Liu, P., Niethammer, M.: Local temperature scaling for probability calibration. In: ICCV (2021)
19. Essen, D.C.V., Uğurbil, K., Auerbach, E.J., Barch, D.M., Behrens, T.E.J., Bucholz, R.D., et al.: The human connectome project: A data acquisition perspective. NeuroImage (2012)
20. Fischl, B., Salat, D.H., Busa, E., Albert, M.S., Dieterich, M., Haselgrove, C., et al.: Whole brain segmentation automated labeling of neuroanatomical structures in the human brain. Neuron (2002)
21. Fonov, V.S., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., et al.: Unbiased average age-appropriate atlases for pediatric studies. Neuroimage (2011)
22. Fonov, V.S., Evans, A.C., McKinstry, R.C., Almli, C.R., Collins, D.: Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. NeuroImage (2009)
23. Fowler, C., Rainey-Smith, S.R., Bird, S., Bomke, J., Bourgeat, P., Brown, B.M., et al.: Fifteen years of the australian imaging, biomarkers and lifestyle (AIBL) study: Progress and observations from 2,359 older adults spanning the spectrum from cognitive normality to Alzheimer's disease. Journal of Alzheimer's disease reports (2021)
24. Henschel, L., Kügler, D., Reuter, M.: FastSurferVINN: Building resolution-independence into deep learning segmentation methods — A solution for highres brain MRI. NeuroImage (2022)
25. Hoffmann, M., Billot, B., Greve, D.N., Iglesias, J.E., Fischl, B., Dalca, A.V.: SynthMorph: Learning contrast-invariant registration without acquired images. IEEE Transactions on Medical Imaging (2020)
26. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: SynthStrip: skull-stripping for any brain image. NeuroImage (2022)
27. Iglesias, J.E., Billot, B., Balbastre, Y., Magdamo, C.G., Arnold, S., Das, S., et al.: SynthSR: A public AI tool to turn heterogeneous clinical brain scans into high-resolution T1-weighted images for 3D morphometry. Science Advances (2023)
28. Iglesias, J.E., Billot, B., Balbastre, Y., Tabari, A., Conklin, J., Alexander, D.C., et al.: Joint super-resolution and synthesis of 1 mm isotropic MP-RAGE volumes from clinical MRI exams with scans of different orientation, resolution and contrast. NeuroImage (2020)
29. Iglesias, J.E., Schleicher, R., Laguna, S., Billot, B., Schaefer, P., McKaig, B., et al.: Quantitative brain morphometry of portable low-field-strength MRI using super-resolution machine learning. Radiology (2022)
30. Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P.M., Alexander, G.E., Harvey, D.J., et al.: The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging (2008)

31. Juntu, J., Sijbers, J., Dyck, D.V., Gielen, J.: Bias field correction for MRI images. In: CORES (2005)
32. Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis (2016)
33. Kapkiç, M., Sağiroğlu, Ş.: Privacy issues in magnetic resonance images. International Journal of Information Security Science (2023)
34. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., et al.: Segment anything. arXiv **abs/2304.02643** (2023)
35. LaMontagne, P.J., Keefe, S.J., Lauren, W., Xiong, C., Grant, E.A., Moulder, K.L., et al.: OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer's disease. Alzheimer's & Dementia (2018)
36. Laso, P., Cerri, S., Sorby-Adams, A., Guo, J., Mateen, F., Goebl, P., et al.: Quantifying white matter hyperintensity and brain volumes in heterogeneous clinical and low-field portable MRI. arXiv **abs/2312.05119** (2023)
37. Leemput, K.V., Maes, F., Vandermeulen, D., Suetens, P.: A unifying framework for partial volume segmentation of brain MR images. IEEE Transactions on Medical Imaging (2003)
38. Li, Z., Tian, L., Mok, T.C.W., Bai, X., Wang, P., Ge, J., et al.: SAMConvex: Fast discrete optimization for CT registration using self-supervised anatomical embedding and correlation pyramid. In: MICCAI (2023)
39. Liu, P., Lee, Y., Aylward, S., Niethammer, M.: Deep decomposition for stochastic normal-abnormal transport. In: CVPR (2022)
40. Liu, P., Tian, L., Zhang, Y., Aylward, S., Lee, Y., Niethammer, M.: Discovering hidden physics behind transport dynamics. In: CVPR (2021)
41. Liu, P., Wang, R., Cao, X., Zhou, Y., Shah, A., Lim, S.N.: Differential motion evolution for fine-grained motion deformation in unsupervised image animation. arXiv **abs/2110.04658** (2021)
42. Milletarì, F., Navab, N., Ahmadi, S.A.: V-Net: Fully convolutional neural networks for volumetric medical image segmentation. 3DV (2016)
43. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., et al.: Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine (2010)
44. Moor, M., Banerjee, O., Abad, Z.F.H., Krumholz, H.M., Leskovec, J., Topol, E.J., Rajpurkar, P.: Foundation models for generalist medical artificial intelligence. Nature (2023)
45. Müller, R.A., Shih, P., Keehn, B., Deyoe, J.R., Leyden, K.M., Shukla, D.: Underconnected, but how? A survey of functional connectivity MRI studies in autism spectrum disorders. Cerebral Cortex (2011)
46. Puonti, O., Iglesias, J.E., Leemput, K.V.: Fast and sequence-adaptive whole-brain segmentation using parametric bayesian modeling. NeuroImage (2016)
47. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: MICCAI (2015)
48. Sheller, M.J., Edwards, B., Reina, G.A., Martin, J., Pati, S., Kotrotsou, A., et al.: Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. Scientific reports (2020)
49. Shen, Z., Feydy, J., Liu, P., Curiale, A.H., San Jose Estepar, R., San Jose Estepar, R., Niethammer, M.: Accurate point cloud registration with robust optimal transport. In: NeurIPS (2021)
50. Singhal, K., Azizi, S., Tu, T., Mahdavi, S., Wei, J., Chung, H.W., et al.: Large language models encode clinical knowledge. Nature (2022)

51. Symms, M.R., Jäger, H.R., Schmierer, K., Yousry, T.A.: A review of structural magnetic resonance neuroimaging. Journal of Neurology, Neurosurgery & Psychiatry (2004)
52. Tanno, R., Worrall, D.E., Kaden, E., Alexander, D.C.: Uncertainty modelling in deep learning for safer neuroimage enhancement: Demonstration in diffusion MRI. NeuroImage (2020)
53. Tian, Q., Bilgiç, B., Fan, Q., Ngamsombat, C., Zaretskaya, N., Fultz, N.E., et al.: Improving in vivo human cerebral cortical surface reconstruction using data-driven super-resolution. Cerebral Cortex (2020)
54. Tu, T., Azizi, S., Driess, D., Schaekermann, M., Amin, M., Chang, P.C., et al.: Towards generalist biomedical AI. arXiv **abs/2307.14334** (2023)
55. de Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Išgum, I.: A deep learning framework for unsupervised affine and deformable image registration. Medical Image Analysis (2019)
56. Wachinger, C., Golland, P., Reuter, M.: BrainPrint: Identifying subjects by their brain. In: MICCAI (2014)
57. Wang, M., Deng, W.: Deep visual domain adaptation: A survey. Neurocomputing (2018)
58. Wang, T.C., Mallya, A., Liu, M.Y.: One-shot free-view neural talking-head synthesis for video conferencing. In: CVPR (2021)
59. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: ACSSC (2003)
60. Weiner, M.W., Veitch, D.P., Aisen, P.S., Beckett, L.A., Cairns, N.J., Green, R.C., et al.: The Alzheimer's disease neuroimaging initiative 3: Continued innovation for clinical trial improvement. Alzheimer's & Dementia (2017)
61. White, T., Blok, E., Calhoun, V.D.: Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed. Human Brain Mapping (2022)
62. Xu, Y., Mo, T., Feng, Q., Zhong, P., Lai, M., Chang, E.I.C.: Deep learning of feature representation with multiple instance learning for medical image analysis. In: ICASSP (2014)
63. Yan, K., Cai, J., Jin, D., Miao, S., Harrison, A.P., Guo, D., et al.: SAM: Self-supervised learning of pixel-wise anatomical embeddings in radiological images. IEEE Transactions on Medical Imaging (2020)
64. Yang, X., Kwitt, R., Niethammer, M.: Quicksilver: Fast predictive image registration – a deep learning approach. NeuroImage (2017)
65. You, C., Zhao, R., Staib, L.H., Duncan, J.S.: Momentum contrastive voxel-wise representation learning for semi-supervised volumetric medical image segmentation. In: MICCAI (2021)
66. Zhou, Y., Chia, M.A., Wagner, S.K., Ayhan, M.S., Williamson, D.J., Struyven, R.R., et al.: A foundation model for generalizable disease detection from retinal images. Nature (2023)