PEPSI: Pathology-Enhanced Pulse-Sequence-Invariant Representations for Brain MRI

Peirong Liu[1] Oula Puonti[1] Annabel Sorby-Adams[1]
W. Taylor Kimberly[1] Juan E. Iglesias[1,2,3]

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

Abstract. Remarkable progress has been made by data-driven machine-learning methods in the analysis of MRI scans. However, most existing MRI analysis approaches are crafted for specific MR pulse sequences (MR contrasts) and usually require nearly isotropic acquisitions. This limits their applicability to the diverse, real-world clinical data, where scans commonly exhibit variations in appearances due to being obtained with varying sequence parameters, resolutions, and orientations – especially in the presence of pathology. In this paper, we propose PEPSI, the first pathology-enhanced, and pulse-sequence-invariant feature representation learning model for brain MRI. PEPSI is trained entirely on synthetic images with a novel pathology encoding strategy, and enables co-training across datasets with diverse pathologies and missing modalities. Despite variations in pathology appearances across different MR pulse sequences or the quality of acquired images (e.g., resolution, orientation, artifacts, etc), PEPSI produces a high-resolution image of reference contrast (MP-RAGE) that captures anatomy, along with an image specifically highlighting the pathology. Our experiments demonstrate PEPSI's remarkable capability for image synthesis compared with the state-of-the-art, contrast-agnostic synthesis models, as it accurately reconstructs anatomical structures while differentiating between pathology and normal tissue. We further illustrate the efficiency and effectiveness of PEPSI features for downstream pathology segmentation on five public datasets covering white matter hyperintensities and stroke lesions. Code is available at https://github.com/peirong26/PEPSI.

1 Introduction

Recent learning-based methods have enabled considerably more rapid and accurate image analysis of brain magnetic resonance imaging (MRI) [15], which provides precise and adjustable soft-tissue contrast via noninvasive, in vivo imaging of the human brain [4]. Nevertheless, the majority of current MRI analysis approaches are tailored to particular MR pulse sequences (MR contrasts), and often rely on nearly isotropic acquisitions. Consequently, sharp declines in performance frequently occur when voxel size and anisotropy increase, or when applied to a contrast that is different from the one used during training [28]. This compromises model generalizability and leads to extra data collection and training efforts when dealing with new datasets. Leveraging synthetic data, recent contrast-agnostic models [16,22,20,15,2,13,18] demonstrate remarkable performance and largely broaden the scope of model applicability to the diverse clinical acquisition protocols. However, these models are confined to the specific tasks they were trained for and cannot be readily adapted to other tasks.

Meanwhile, task-agnostic foundation models [3,1] in general computer vision and natural language processing have experienced notable success, driven by the fast growth of large-scale datasets [5,8,17]. Nonetheless, the development of foundation models in medical imaging has been hindered by the lack of large-scale datasets (in many domains), variations in acquisition protocols and processing pipelines, and privacy constraints. MONAI [7] provides pre-trained models for diverse tasks, but they generally are highly task-oriented and contrast-sensitive. Zhou et al. [30] proposed a medical foundation model, which is specifically designed for the detection of eye and systemic health conditions from retinal scans, yet this model is limited to the modalities of color fundus photography and optical coherence tomography. AI generalist systems [24,26,27] have shown superiority in biomedical tasks (e.g., visual question answering, image classification, radiology report generation and summarizing), but mostly within the vision-language context. CIFL [9] was designed for task-agnostic feature representations, yet it has only been demonstrated in 2D, and exclusively relies on contrastive learning, insufficient in surpassing task-specific models in downstream applications [21]. Recently, Liu et al. [21] proposed Brain-ID, which extracts contrast-agnostic features for brain MRI, and achieves state-of-the-art performance in various fundamental medical imaging tasks including reconstruction, segmentation, and super-resolution. However, Brain-ID exclusively focuses on healthy-appearing anatomy and lacks the capacity to model pathologies (Fig. 4).

In this paper, we introduce `PEPSI`, the first pulse-sequence-invariant feature representation learning approach specifically designed to emphasize pathology. `PEPSI` is trained on synthetic data encoded with pathology, and can be directly applied to real images featuring various types of pathology.

1) We introduce a data generator that synthesizes images incorporating augmented pathologies across *any* combination of deformation, pulse sequence, resolution, orientation, artifacts, etc., thus circumventing the limitations of real data, which are often confined to the acquired pulse sequence (Fig. 1).
2) We design a feature learning framework guided by MP-RAGE and FLAIR scans, which balances anatomy and pathology. Furthermore, `PEPSI` bridges the gaps of pathologies across datasets via our proposed implicit pathology supervision, and enables co-training across datasets with different pathology types and potentially missing modalities (Sec. 2.2).
3) We conduct comprehensive evaluations on image synthesis and pathology segmentation. `PEPSI` exhibits *(i)* a remarkable capability to synthesize images with missing modalities while simultaneously capturing various pathologies (Fig. 4); *(ii)* superior efficiency and effectiveness on downstream pathology segmentation across five public datasets, covering modalities of T1w and FLAIR, with white matter hyperintensity (WMH) and stroke lesions (Tab. 2).

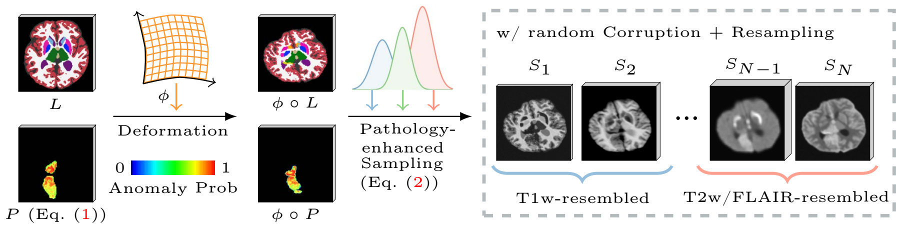

Fig. 1. PEPSI's on-the-fly generator uses 3D anatomy labels (L) and anomaly probabilities (P) to generate training data with diverse deformations, contrasts, and corruptions – enhanced by varying intensity profiles in pathological regions (Sec. 2.1).

2 Approach

Sourcing large-scale datasets with high-quality and diverse contrasts for brain MRI is challenging. Recent works [2,14,15,21] proposed to utilize anatomy labels to simulate data, yet their data generators are *solely* based on anatomy and lack prior information on potential pathologies. Instead, we synthesize data that *emphasizes* pathologies (Sec. 2.1), to encourage the model to *distinguish* between normal and abnormal regions in the features (Sec. 2.2), thereby facilitating the transmission of valuable information for downstream pathology segmentations.

2.1 Generating Pathology-encoded Training Data

PEPSI leverages neuroanatomical labels and pathology segmentation to generate contrast-diverse data while *simultaneously* emphasizing pathology.

Anomaly Probabilities: We construct a proxy for anomaly maps (P) using *a priori* knowledge of an image's expected appearance conditioned on its MR contrast, where pathology is typically darker in T1w and brighter in T2w/FLAIR:

$$P(x) = \begin{cases} 0\,, & x \notin \Omega_P \\ 1 - (I(x) - I_{\min})/(I_{\max} - I_{\min})\,, & x \in \Omega_P\,,\ I \in \{\text{T1w}\} \\ (I(x) - I_{\min})/(I_{\max} - I_{\min})\,, & x \in \Omega_P\,,\ I \in \{\text{T2w, FLAIR}\} \end{cases} \tag{1}$$

where Ω_P refers to the pathological region, $I_{\max}$ ($I_{\min}$) is the regional maximum (minimum) image intensities: $I_{\max} = \max_{x \in \Omega_P} I(x)$, $I_{\min} = \min_{x \in \Omega_P} I(x)$.

Pathology-encoded Contrast: To generate images with complex brain structures, we leverage anatomy labels following [21]. As shown in Fig. 1, a random deformation field (ϕ) is first generated, comprising linear and non-linear transformations [16,21]. After the anatomy labels (L) and anomaly probabilities (P) are deformed by ϕ, we generate the pathology-encoded images via two steps:
(i) "Anomaly-free" image (S_0): We begin with randomly sampling intensities on the brain anatomy labels, where the regional intensities are generated by independently sampling a Gaussian distribution for each labeled region [21].

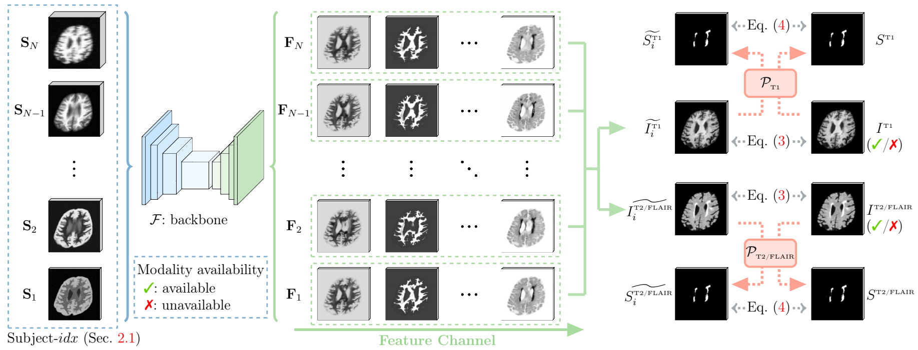

Fig. 2. `PEPSI`'s pathology-enhanced, contrast-agnostic training overview (Sec. 2.2).

(ii) Pathology enhancement: We incorporate the anomaly probabilities into the "anomaly-free" image (S_0) to produce a pathology-encoded image (S) – again, using *a priori* knowledge of the modality. This is conditioned on the direction of intensities from white to gray matter in S_0: $S(x) = S_0(x) + \Delta S(x) * p(x)$,

$$\text{s.t. } \Delta S(x) \sim \begin{cases} \{0\}\,, & x \notin \Omega_{\phi \circ P} \\ \mathcal{N}(-\mu_\text{w}/2,\, \mu_\text{w}/2)\,, & x \in \Omega_{\phi \circ P}\,, \mu_\text{w} > \mu_\text{g} \\ \mathcal{N}(\mu_\text{w}/2,\, \mu_\text{w}/2)\,, & x \in \Omega_{\phi \circ P}\,, \mu_\text{w} \leq \mu_\text{g} \end{cases} \tag{2}$$

μ_w (μ_g) is the mean value of white (gray) matter intensities in S_0. A higher μ_w resembles T1w, where pathologies appear darker; A lower μ_w resembles T2w or FLAIR, where pathologies are typically brighter. (See the dashed box in Fig. 1.)

The pathology-encoded images (S) further undergo a corruption pipeline [15] (Fig. 1), which includes a model of partial voluming [2], and introduces various resolutions, noises and scanning artifacts commonly found in clinical protocols.

2.2 Representing across Contrasts, Pathologies, Datasets

Here we present `PEPSI`'s training framework, which learns to emphasize anomalies and facilitates co-training across datasets with different types of pathology.

Input: We adopt the "mild-to-severe" intra-subject sampling strategy in [21], which maximizes intra-subject variance to enhance feature robustness. Samples generated within a mini-batch are from the same subject, yet exhibit varying contrasts, corruptions, and pathology intensities, enriching the learning space (Fig. 2).

Dual Guidance Balancing Anatomy and Pathology: We aim to obtain robust, contrast-agnostic feature representations that capture the distinctive anatomy of each subject while effectively distinguishing between pathology and normal tissue. MP-RAGE is the standard T1w MR contrast to delineate anatomical structures in research, but it is insufficient to differentiate many types of anomalies from normal tissue. FLAIR MRI, on the other hand, highlights areas

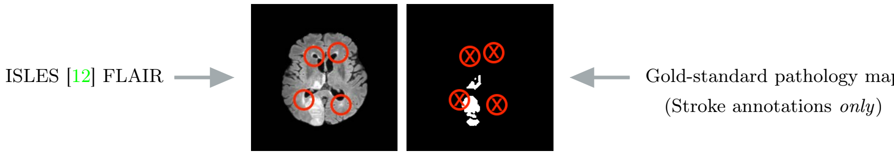

Fig. 3. Left: an axial slice of a FLAIR from ISLES [12], WMH marked in red. Right: its gold-standard lesion segmentation, which only includes stroke lesions (no WMH).

of T2 prolongation as bright while suppressing cerebrospinal fluid (CSF), providing clear visibility of lesions in proximity to CSF [11] – but provides worse contrast than MP-RAGE in normal anatomy. PEPSI resorts to both MP-RAGE and FLAIR as learning targets, to concurrently capture normal anatomy and pathology. (Fig. 4 compares the performance of dual-guidance and single-guidance.)

As shown in Fig. 2, the input mini-batch of intra-subject pathology-encoded samples, $\{S_1, \ldots, S_N\}$, are mapped to their corresponding feature space by a backbone ($\mathcal{F}$), $\{\mathbf{F}_1, \ldots, \mathbf{F}_N\}$. Two linear activation layers are followed to synthesize the anatomy and pathology images. The synthesis loss is obtained by collecting the reconstruction errors of all samples in the current mini-batch:

$$\begin{aligned}\mathcal{L}_{\text{synth}} &= \alpha \mathcal{L}_{I_{\text{T1}}} + \beta \mathcal{L}_{I_{\text{T2/FLAIR}}} \qquad\qquad (\alpha,\, \beta \in \{0,\, 1\},\, \lambda \in R^+) \qquad (3)\\ &= \alpha \sum\nolimits_i^N \, |\widetilde{I_i^{\text{T1}}} - I^{\text{T1}}| + \lambda\, |\nabla \widetilde{I_i^{\text{T1}}} - \nabla I^{\text{T1}}| \\ &\quad + \beta \sum\nolimits_i^N \, |\widetilde{I_i^{\text{T2/FLAIR}}} - I^{\text{T2/FLAIR}}| + \lambda\, |\nabla \widetilde{I_i^{\text{T2/FLAIR}}} - \nabla I^{\text{T2/FLAIR}}|,\end{aligned}$$

where I ($\tilde{I}$) is the ground truth (predicted) image, α (β) denote the availability of ground truth I^{T1} ($I^{\text{T2/FLAIR}}$), λ is the weight of reconstruction gradient loss [21].

Implicit Pathology Supervision for Multi-pathology/dataset Training: Co-training across datasets broadens the model's exposure to various types of pathology, but also presents inherent challenges – notably, difficulty to accurately synthesize *abnormal* regions in the *missing* modality, particularly for smaller datasets (e.g., "PEPSI (No-Seg)" in Fig. 4). Direct supervision on pathology segmentations forces the model to pay more attention to anomalies, but could potentially result in conflicts during co-training due to the *non-exhaustive* pathology annotations across datasets (e.g., "PEPSI (Dir-Seg)" in Fig. 4) – The above figure shows a FLAIR image from ISLES [12] stroke dataset, despite the acquired FLAIR image clearly indicating WMH (circled in red), their gold-standard pathology segmentation *only* provides/annotates areas of stroke lesions.

Here we propose an indirect pathology supervision approach. Specifically, for each output modality (i.e., MP-RAGE and FLAIR), we employ a "third-party", real-image-supervised pathology segmentation model as a reference, to encourage the pathology estimated from the predicted and ground truth images to align, without imposing strict supervision from the gold-standard pathology maps. As depicted in Fig. 2, we pass all intra-subject training samples through

Table 1. Quantitative comparisons in anatomy and pathology image synthesis among `PEPSI`, its variants, and the state-of-the-art contrast-agnostic synthesis models. The proposed `PEPSI` *(i)* outperforms all the other models, especially on single-modality datasets; and *(ii)* preserves its high performance even for cross-modality synthesis.

Dataset (# of train/test)	MR Contrast Input	MR Contrast Output	Metric	SynthSR [15]	Brain-ID [21]	PEPSI (SG-T1w)	PEPSI (SG-FLAIR)	PEPSI (No-Seg)	PEPSI (Dir-Seg)	PEPSI (Proposed)
ATLAS [19] (590/65)	T1w	T1w	L1 (↓)	0.067	0.65	0.69	-	0.052	0.074	**0.036**
			PSNR (↑)	16.90	17.91	16.54	-	18.46	16.01	**21.69**
			SSIM (↑)	0.804	0.833	0.845	-	0.861	0.831	**0.897**
ISLES [12] (137/15)	FLAIR	FLAIR	L1 (↓)	-	-	-	0.022	0.018	0.021	**0.016**
			PSNR (↑)	-	-	-	23.87	25.34	24.02	**26.03**
			SSIM (↑)	-	-	-	0.962	0.942	0.926	**0.969**
ADNI3 [29] (298/33)	T1w	T1w	L1 (↓)	0.023	0.021	0.025	-	0.022	0.022	**0.020**
			PSNR (↑)	23.51	24.42	24.44	-	24.01	23.37	**26.67**
			SSIM (↑)	0.901	0.899	0.930	-	0.932	0.931	**0.935**
		FLAIR	L1 (↓)	-	-	-	0.043	0.392	0.396	**0.036**
			PSNR (↑)	-	-	-	18.87	19.64	19.58	**21.40**
			SSIM (↑)	-	-	-	0.900	0.901	0.894	**0.911**
	FLAIR	T1w	L1 (↓)	0.027	0.026	0.027	-	0.027	0.029	**0.023**
			PSNR (↑)	23.25	23.74	23.96	-	23.50	23.61	**25.62**
			SSIM (↑)	0.906	0.879	0.916	-	0.919	0.915	**0.929**
		FLAIR	L1 (↓)	-	-	-	0.044	0.0396	0.041	**0.034**
			PSNR (↑)	-	-	-	18.65	19.66	19.31	**21.77**
			SSIM (↑)	-	-	-	0.911	0.910	0.904	**0.914**

the frozen, reference pathology segmentation models ($\mathcal{P}_{\mathrm{T1}}$, $\mathcal{P}_{\mathrm{T2/FLAIR}}$). The implicit pathology loss is computed based on the segmentation errors between the estimated pathology maps from the synthesized and ground truth images:

$$\begin{aligned}\mathcal{L}_{\mathrm{pathol}} &= \alpha \mathcal{L}_{S_{\mathrm{T1}}} + \beta \mathcal{L}_{S_{\mathrm{T2/FLAIR}}} \qquad (\alpha,\, \beta \in \{0,\, 1\}) \\ &= \alpha \sum\nolimits_i^N \, \mathcal{L}_{\mathrm{seg}}(\widetilde{S_i^{\mathrm{T1}}},\, S^{\mathrm{T1}}) + \beta \sum\nolimits_i^N \, \mathcal{L}_{\mathrm{seg}}(\widetilde{S_i^{\mathrm{T2/FLAIR}}},\, S^{\mathrm{T2/FLAIR}})\,.\end{aligned} \tag{4}$$

$\mathcal{L}_{\mathrm{seg}}$ is the segmentation loss consisting of soft dice and cross-entropy loss [2]. S ($\tilde{S}$) denotes the third-party-referenced (predicted) pathology[1]. Therefore, the overall training object writes $\mathcal{L} = \mathcal{L}_{\mathrm{T1}} + \omega \mathcal{L}_{\mathrm{T2/FLAIR}}$, $\omega \in \mathcal{R}^+$.

3 Experiments

We demonstrate the effectiveness of `PEPSI` from two perspectives: *(i) Image synthesis* — estimating both anatomy and pathology images, with potentially missing modalities (Sec. 3.1); *(ii) Pathology segmentation* — fine-tuning `PEPSI` on individual datasets for segmenting a specific type of pathology (Sec. 3.2).

Datasets: To cover a broader range of anatomies and pathologies, we train `PEPSI` on 1025 subjects from (# of train/test cases): *(i)* ADNI3 [29] (298/33), with 1 *mm* isotropic T1w and FLAIR pairs with WMH; *(ii)* ATLAS [19] (590/65), with *only* T1w and manually segmented stroke lesion for subacute/chronic stroke

[1] We train a segmentation model using data with *minimal corruption*, since it shall work well *only if* the inputs are of high quality — It would be uninformative if the segmentation network provides accurate labels for images of any quality.

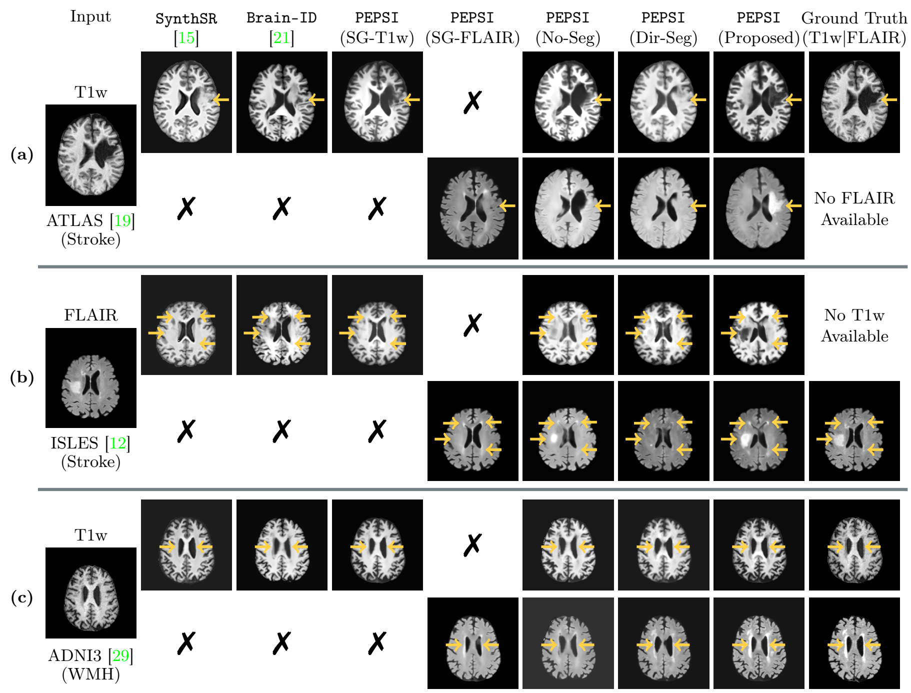

Fig. 4. Qualitative comparisons on T1w and FLAIR synthesis (↔ highlights the ground truth regions of pathology). Rows (columns) refer to the datasets (compared methods).

patients; *(iii)* ISLES [12] (137/15), with *only* FLAIR and stroke lesion segmentation for acute/subacute stroke patients. For pathology segmentation, we also test on ISBI2015 [6] and MSSEG2016 [10], comprising 21 and 15 WMH patients.

Metrics: For image synthesis, we use `L1` distance, `PSNR`, and `SSIM` (structural similarity) [23]. For pathology segmentation, we use `Dice` scores [2].

Models: We compare `PEPSI` with the state-of-the-art contrast-agnostic synthesis methods, `SynthSR` [15] and `Brain-ID` [21]. We also evaluate `PEPSI`'s variants: *(i-ii)* SG-T1w/FLAIR: single-guidance from MR-RAGE/FLAIR; *(iii-iv)* No/Dir-Seg: No/direct supervision from gold-standard pathology segmentations.

Implementation Details: As a general feature representation model, `PEPSI` can use any backbone to extract features. For fairer comparison, we adopt the same five-level 3D UNet [25] as utilized in state-of-the-art models [15,21] we compare with. Two linear layers are followed for anatomy and pathology image synthesis (Sec. 2.2). The synthetic pathology-encoded data is of size 128^3 (Sec. 2.1), with batch size as 4. We use AdamW optimizer, with a learning rate of 10^{-4} for the first 160,000 iterations and 10^{-5} until 240,000 iterations. We set $\lambda = 1$ in Eq. (3), and $\omega = 0.1$ in Sec. 2.2 for 100,000 iterations, and 1 afterward. The training took ≈ 5 days on one NVIDIA RTX8000 GPU.

Table 2. Average `Dice` scores (↑) for pathology segmentation, w/o or w/ `PEPSI` pre-trained features. (Numbers in the parentheses denote the convergence/testing epochs (↓); We directly test on ISBI2015 and MSSEG2016 using models trained from ADNI3.)

Model	ATLAS (Stroke)	ISLES (Stroke)	ADNI3 (WMH)		ISBI2015 (WMH)		MSSEG2016 (WMH)	
	T1w	FLAIR	T1w	FLAIR	T1w	FLAIR	T1w	FLAIR
w/o `PEPSI`	0.49 ± 0.14 (2500)	0.35 ± 0.13 (2000)	0.50 ± 0.15 (1600)	0.67 ± 0.13 (1500)	0.21 ± 0.05 (1600)	0.39 ± 0.15 (1500)	0.24 ± 0.09 (1600)	0.31 ± 0.10 (1500)
w/ `PEPSI`	$\mathbf{0.71} \pm 0.22$ (**1000**)	$\mathbf{0.62} \pm 0.27$ (**500**)	$\mathbf{0.69} \pm 0.12$ (**800**)	$\mathbf{0.75} \pm 0.10$ (**500**)	$\mathbf{0.34} \pm 0.06$ (**800**)	$\mathbf{0.57} \pm 0.15$ (**500**)	$\mathbf{0.38} \pm 0.10$ (**800**)	$\mathbf{0.45} \pm 0.11$ (**500**)

3.1 Anatomy and Pathology Image Synthesis

As shown in Tab. 1, `PEPSI` achieves the best performance in synthesizing T1w and FLAIR, across all datasets and pathologies. Notably, `PEPSI` exhibits superiority on single-modality datasets (ATLAS [19], ISLES [12]), and further demonstrates strong *robustness* against contrasts. For example, it maintains consistent scores for T1w synthesis on ADNI3 [29], regardless of the input modality.

Thanks to the co-training and pathology-enhanced, contrast-agnostic learning, `PEPSI` can synthesize images that are *not* present in the original datasets. Fig. 4-(a): `PEPSI` successfully synthesizes T1w and pathology-enhanced images based on T1w from ATLAS [19], for which ground truth FLAIR is not available. Remarkably, other models either *cannot* estimate pathology-enhanced images, or struggle to *accurately* capture and highlight (brighten) the areas of pathology. Fig. 4-(b): ISLES [12] only provides FLAIR and annotations for stroke lesions, yet `PEPSI`: *(i)* accurately synthesizes T1w images with appropriately *darkened* pathology regions inferred from the FLAIR input, and *(ii)* is not constrained to the stroke lesions manually annotated by ISLES, but instead, captures (*brightens*) all pathological regions including *both* stroke lesions and WMH.

3.2 Pathology Segmentation

In Sec. 3.1, we validate `PEPSI`'s superiority in synthesizing pathology-enhanced images under various contrasts, providing voxel-level information that is *not* confined to particular pathology types, but contains comprehensive information on anomalies. We further illustrate the *efficiency* and *effectiveness* of `PEPSI` features for downstream pathology segmentations that target a specific pathology.

To this end, we compare the following two models trained on each dataset and contrast, *(i)* starting from random weights (w/o `PEPSI`), and *(ii)* fine-tuned from `PEPSI` pre-trained weights (w/ `PEPSI`). For ATLAS [19], ISLES [12], and ADNI3 [29], both models are trained and tested on their respective training and testing sets. Since ISBI2015 [6] and MSSEG2016 [10] datasets contain only 21 and 15 WMH cases, respectively, we directly evaluate the trained models from ADNI3 (WMH) [29] on all available cases in these datasets. Note that although `PEPSI` has undergone pre-training on synthetic data using anatomy labels and pathology probability maps from the training sets of ATLAS [19] and ISLES [12] (Sec. 3), it has *not* been exposed to any real image during the pre-training stage.

As shown in Tab. 2, utilizing PEPSI's pre-trained features largely reduces the convergence time (by $\approx 60\%$ on average) – and, more importantly, it yields higher Dice scores compared with training from scratch (i.e., w/o PEPSI) on all testing pathologies, contrasts and datasets. Furthermore, when directly tested on the two small datasets (ISBI2015 and MSSEG2016), PEPSI exhibits superior generalizability compared to models trained without PEPSI pre-trained features.

4 Conclusion

We introduced PEPSI, the first pathology-enhanced, contrast-agnostic feature representation learning approach for brain MRI. Trained on synthetic data with diverse contrasts and anomalies, PEPSI exhibits remarkable robustness and accuracy beyond manual annotations of a specific pathology, regardless of MR contrasts. We demonstrated PEPSI's performance on image synthesis, covering T1w and FLAIR with stroke lesions and WMH, and further showcased its efficiency and effectiveness for downstream pathology segmentation on five public datasets. We believe PEPSI will pave the way for the exciting future of contrast-agnostic pathology representations for heterogeneous, real-world brain MRI – enabling studies of diverse brain diseases with large clinical MRI datasets.

Acknowledgments. Primarily supported by NIH 1RF1AG080371. Additional support from NIH 1RF1AG080371, 1R01EB031114, 1R01AG070988, 1RF1MH123195. ASA is funded by the American Heart Association Postdoctoral Fellowship and the Fulbright Commission. OP is supported by Lundbeckfonden R360–2021–395.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. 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) 2
2. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Leemput, K.V., Fischl, B.R., et al.: SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. Medical Image Analysis (2021) 2, 3, 4, 6, 7
3. 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) 2
4. 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) 1
5. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., et al.: Language models are few-shot learners. In: NeurIPS (2020) 2
6. Carass, A., Roy, S., Jog, A., Cuzzocreo, J.L., Magrath, E., Gherman, A., et al.: Longitudinal multiple sclerosis lesion segmentation: resource and challenge. NeuroImage (2017) 7, 8

7. 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) 2
8. Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., et al.: PaLM: Scaling language modeling with pathways. JMLR (2022) 2
9. Chua, Y.Z.R., Dalca, A.V.: Contrast invariant feature representations for segmentation and registration of medical images. In: MIDL (2023) 2
10. Commowick, O., Kain, M., Casey, R., Ameli, R., Ferré, J.C., Kerbrat, A., et al.: Multiple sclerosis lesions segmentation from multiple experts: The MICCAI 2016 challenge dataset. Neuroimage (2021) 7, 8
11. Hajnal, J.V., Bryant, D.J., Kasuboski, L., Pattany, P.M., De Coene, B., Lewis, P.D., et al.: Use of fluid attenuated inversion recovery (FLAIR) pulse sequences in MRI of the brain. Journal of computer assisted tomography (1992) 5
12. Hernandez Petzsche, M.R., de la Rosa, E., Hanning, U., Wiest, R., Valenzuela, W., Reyes, M., et al.: ISLES 2022: A multi-center magnetic resonance imaging stroke lesion segmentation dataset. Scientific data (2022) 5, 6, 7, 8
13. Hoffmann, M., Billot, B., Greve, D.N., Iglesias, J.E., Fischl, B.R., Dalca, A.V.: SynthMorph: Learning contrast-invariant registration without acquired images. IEEE Transactions on Medical Imaging (2020) 2
14. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B.R., Hoffmann, M.: SynthStrip: skull-stripping for any brain image. NeuroImage (2022) 3
15. 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) 1, 2, 3, 4, 6, 7
16. 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) 2, 3
17. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., et al.: Segment anything. arXiv **abs/2304.02643** (2023) 2
18. 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) 2
19. Liew, S.L., Anglin, J.M., Banks, N.W., Sondag, M., Ito, K.L., Kim, H., et al.: A large, open source dataset of stroke anatomical brain images and manual lesion segmentations. Scientific data (2018) 6, 7, 8
20. Liu, P., Lee, Y., Aylward, S., Niethammer, M.: Deep decomposition for stochastic normal-abnormal transport. In: CVPR (2022) 2
21. Liu, P., Puonti, O., Hu, X., Alexander, D.C., Iglesias, J.E.: Brain-ID: Learning robust feature representations for brain imaging. In: ECCV (2024) 2, 3, 4, 5, 6, 7
22. Liu, P., Tian, L., Zhang, Y., Aylward, S., Lee, Y., Niethammer, M.: Discovering hidden physics behind transport dynamics. In: CVPR (2021) 2
23. 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) 7
24. 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) 2
25. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: MICCAI (2015) 7

26. Singhal, K., Azizi, S., Tu, T., Mahdavi, S., Wei, J., Chung, H.W., et al.: Large language models encode clinical knowledge. Nature (2022) 2
27. Tu, T., Azizi, S., Driess, D., Schaekermann, M., Amin, M., Chang, P.C., et al.: Towards generalist biomedical AI. arXiv **abs/2307.14334** (2023) 2
28. Wang, M., Deng, W.: Deep visual domain adaptation: A survey. Neurocomputing (2018) 1
29. 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) 6, 7, 8
30. 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) 2