Symmetry Awareness Encoded Deep Learning Framework for Brain Imaging Analysis

Yang Ma[1,2], Dongang Wang[2,5], Peilin Liu[2,3], Lynette Masters[4], Michael Barnett[2,5], Weidong Cai[1], and Chenyu Wang[2,5]

[1] School of Computer Science, University of Sydney, NSW 2008, Australia
[2] Brain and Mind Centre, University of Sydney, NSW 2050, Australia
[3] School of Mathematics and Statistics, University of Sydney, NSW 2050, Australia
[4] I-MED Radiology Network, NSW 2050, Australia
[5] Sydney Neuroimaging Analysis Centre, NSW 2050, Australia
chenyu.wang@sydney.edu.au

Abstract. The heterogeneity of neurological conditions, ranging from structural anomalies to functional impairments, presents a significant challenge in medical imaging analysis tasks. Moreover, the limited availability of well-annotated datasets constrains the development of robust analysis models. Against this backdrop, this study introduces a novel approach leveraging the inherent anatomical symmetrical features of the human brain to enhance the subsequent detection and segmentation analysis for brain diseases. A novel Symmetry-Aware Cross-Attention (SACA) module is proposed to encode symmetrical features of left and right hemispheres, and a proxy task to detect symmetrical features as the Symmetry-Aware Head (SAH) is proposed, which guides the pretraining of the whole network on a vast 3D brain imaging dataset comprising both healthy and diseased brain images across various MRI and CT. Through meticulous experimentation on downstream tasks, including both classification and segmentation for brain diseases, our model demonstrates superior performance over state-of-the-art methodologies, particularly highlighting the significance of symmetry-aware learning. Our findings advocate for the effectiveness of incorporating symmetry awareness into pretraining and set a new benchmark for medical imaging analysis, promising significant strides toward accurate and efficient diagnostic processes. Code is available at https://github.com/bitMyron/saswin.

Keywords: Symmetry-Aware Cross-Attention (SACA) · Self-Supervised Learning · Neuroimaging Diagnosis

1 Introduction

In many neurological and psychiatric conditions, the symmetry nature of the left and right brain hemispheres plays a pivotal role in brain disease diagnosis and monitoring, often acting as a precursor to various disorders. The measurement of pathological structural and functional asymmetries could also be used in managing disease progression and the effectiveness of treatment [18]. The significance of

these asymmetrical alterations is especially marked in diseases like Alzheimer's disease (AD) [14] and schizophrenia [22], where hemispheric symmetry is also considered as one of the most indicative biomarkers by radiologists. Figure 1 showcases instances of pathologies affecting the brain's symmetry.

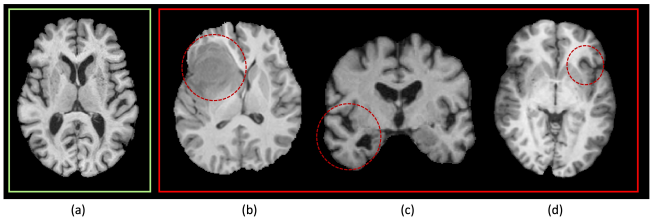

Fig. 1: Brain T1 images: (a) represents a healthy case, displaying symmetrical shape, structure, and image intensity. (b)-(d) depict patients with a brain tumor, Alzheimer's Disease, and focal epilepsy, respectively, and the noticeable structural asymmetry parts are circled out.

Leveraging the clinical experience and identifying such asymmetrical features in neuroimaging analysis tasks can further enhance the efficacy of deep learning tools used in triage, annotation, and diagnosis. Certain approaches attempt to identify asymmetries by subtracting the L/R flipped image from the original and using the resulting subtraction as input for deep neural networks [12]. Other strategies involve segmenting the regions of interest on the original and flipped images separately, employing the pair of segmentation masks to compute an additional loss to improve model performance [13]. However, hemispheric symmetrical features were only used for pre-processing or data augmentation rather than being encoded within the network structure, which limits the generalization ability to extend to the general neuroimaging analysis process.

To better leverage symmetrical features, we introduce an innovative Symmetry-Aware Cross-Attention (SACA) module, advancing the attention mechanism's application in neuroimaging analysis. Our module enhances the transformer blocks' capabilities by implementing cross-attention mechanisms between the original input and its symmetrical counterpart, compelling the network to focus on symmetrical comparisons and semantic feature encoding. This approach mimics the expert diagnostic strategy, leveraging self-supervised and supervised training processes to improve the model's understanding of brain anatomy and its inherent symmetries.

The SACA module can be pretrained with a substantial amount of data based on a self-supervised learning framework to be further used for downstream tasks, including disease diagnosis, brain structure, and lesion segmentation. The advancement of self-supervised learning has been illustrated in previous work [1], which involve generating pre-text tasks, including inpainting [10], rotation [9] and contrastive coding [2], to embed intrinsic features. Besides these tasks widely

used in self-supervised learning methods, we introduce a Symmetry-Aware Head (SAH) to enhance the pretraining process by focusing on the contrastive symmetrical features of the left and right hemispheres. Since they are entangled with the patient status (i.e., healthy or disease), the hemispheric symmetrical features are utilized in computing the symmetry-aware loss.

To fulfill the pretraining, we have compiled an extensive dataset comprising brain MRI and CT images from both healthy subjects (assumed to be symmetrical) and diseased subjects (assumed with asymmetrical features). Empirical experiments of our approach, through extensive pretraining on brain CT and MRI datasets and subsequent evaluations on various classification and segmentation tasks, underscore a notable improvement over existing baseline models.

In general, our primary contributions are outlined as follows:

- We introduce a novel Symmetry-Aware Cross-Attention (SACA) module, which leverages a cross-attention mechanism to analyze the relationship between an image and its symmetrical counterpart, facilitating a deeper understanding of brain anatomy and its inherent symmetries.
- The proposed network is pretrained on a symmetry-aware self-supervised process and can be further applied in real-world clinical datasets and analysis tasks, including classification and segmentation for multiple modalities.
- A vast dataset spanning MRI and CT modalities has been curated, and our approach has undergone extensive testing and demonstrated state-of-the-art performance on diagnosis and segmentation tasks.

2 Methodology

In this study, we leverage a substantial corpus of 3D brain imaging data, denoted as $\mathcal{I}^0 = \{\mathbf{I}_n\}$. Each image is associated with a binary label $\mathcal{Y}^0 = \{0, 1\}$ indicating the health status of the subject (healthy or diseased), where $y_n^0 \in \mathcal{Y}^0$. Our objective is to train a model to encapsulate structural insights into the human brain. The comprehensive pipeline is depicted in Figure 2.

2.1 Preprocssing

A midway registration pipeline was introduced to standardize the input image to position the brain within a standardized 3D coordinate framework, ensuring that flipping the image volume along its sagittal axis could generally map the left and right cerebral hemispheres correctly. Given an original input image $\mathbf{I}$, we first generate its mirror image $\mathbf{I}'$ by flipping $\mathbf{I}$ across the vertical plane of the image. We then obtain the affine transformation $\mathcal{T} : \mathbf{I} \rightarrow \mathbf{I}'$ with the widely used linear registration tool FMRIB's Linear Image Registration Tool (FLIRT) [17]. Based on the generated transformation matrix, the midway points of all transformation operations can be further calculated and formed the midway transformation $\mathcal{T}_{1/2}$. The midway transformation is applied to $\mathbf{I}$, resulting in $\bar{\mathbf{I}}$, whose sagittal plane is perfectly aligned with respect to the vertical plane of the image volume. By

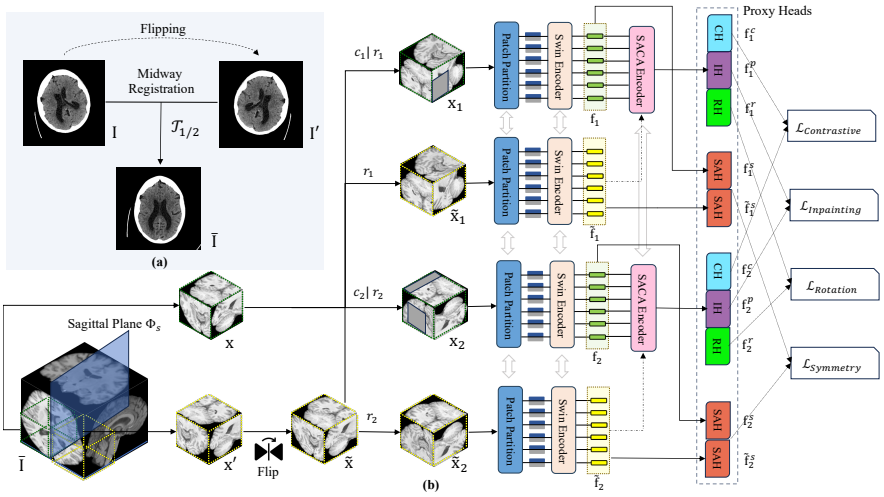

Fig. 2: (a) The process of midway registration to align input images; (b) The framework applied to input sub-volumes and their sagittal symmetrical counterparts. This framework is augmented by the Symmetry-Aware Cross-Attention (SACA) module and is optimized through a composite loss function that integrates inpainting, rotation, contrastive, and symmetry losses.

the midway registration, the simple flip of the input image ensures the mirroring of the left and right hemispheres, assisting the model in learning and utilizing the inherent hemispheric symmetry of the human brain.

After the midway registration, random sub-volumes, denoted as $\mathcal{X} = \{\mathbf{x}_n\}$, are extracted as inputs to networks from the collection of all aligned images $\bar{\mathcal{I}} = \{\bar{\mathbf{I}}_n\}$. Since the input images are standardized by the midway registration, for each patch $\mathbf{x} \in \mathbb{R}^{H \times W \times D}$, the mirrored counterpart $\mathbf{x}'$ with respect to the sagittal plane Φ_s can be identified and can be flipped to obtain $\tilde{\mathbf{x}}$. Note that the generated $\tilde{\mathbf{x}}$ is expected to closely resemble the original patch $\mathbf{x}$, barring minor structural discrepancies inherent to the intrinsic hemispheric symmetry of the human brain. Both original patch $\mathbf{x}$'s and their flipped symmetrical counterparts $\tilde{\mathbf{x}}$'s are used as the input to the encoder network as sub-volumes.

As illustrated in Figure 2, two distinct rotations $R = \{r_1, r_2\}$ are conducted on both sub-volumes and two random inner cutouts $C = \{c_1, c_2\}$ are conducted on original sub-volume $\mathbf{x}$, resulting in augmented sub-volumes $\{\mathbf{x}_1, \mathbf{x}_2, \tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2\}$. The augmented sub-volumes are segmented into patches of the size (h, w, d), ending up with a sequence of the length $\frac{H}{h} \times \frac{W}{w} \times \frac{D}{d}$. These patches are processed through an embedding network and subsequently passed into a 3D Sliding-window (Swin) Transformer architecture [20], which consists of four stages and eight transformer layers. This results in a reduced sequence of the length $\frac{H}{16h} \times \frac{W}{16w} \times \frac{D}{16d}$ with the embedding feature dimension 768.

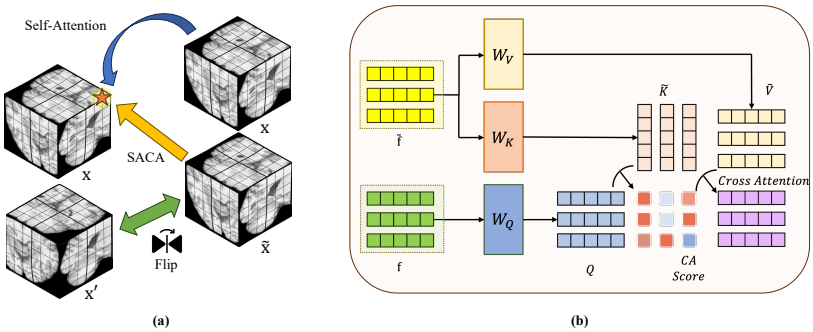

Fig. 3: Symmetry-Aware Cross-Attention (SACA) Module: (a) The computation of attention for each token within the SACA module; (b) Cross-attention calculation of SACA module structure.

2.2 Symmetry-Aware Cross-Attention (SACA) Module

Following the Swin Transformer Encoder [11], we introduce a Symmetry-Aware Cross-Attention (SACA) module to foster deep symmetrical feature alignment and referencing. The SACA module processes input embedding sequences $\mathbf{f}$, applying self-attention mechanisms to $\mathbf{f}$ and cross-attention to the flipped symmetrical counterpart $\tilde{\mathbf{f}}$. This dual attention strategy, as depicted in Figure 3, is intended to exchange information between the raw sub-volume and the flipped symmetrical counterpart.

The cross-attention mechanism within this module is represented as follows:

$$\text{SACA}(\mathbf{f}, \tilde{\mathbf{f}}) = \text{Softmax}\left(\frac{Q\tilde{K}^T}{\sqrt{D}}\right)\tilde{V}, \text{where} \begin{cases} Q = \mathbf{f}W_Q \\ \tilde{K} = \tilde{\mathbf{f}}W_K \\ \tilde{V} = \tilde{\mathbf{f}}W_V \end{cases}. \tag{1}$$

Q represents the queries generated from the features of the original patch $\mathbf{f}$, and $\tilde{K}$ and $\tilde{V}$ are keys and values generated from the counterpart patch $\tilde{\mathbf{f}}$, with D denoting the dimensionality of the query and key vectors. The matrices W are learnable weights for the query, key, and value transformations within the SACA module. With our novel SACA module, the contrastive information from the original patch and its flipped symmetrical counterpart could be utilized, and the level of symmetry of the patch could be learned along the process, enabling the network to recognize symmetrical features of input brain images.

2.3 Symmetry-Aware Pretraining Proxies

Upon obtaining the encoded features from the SACA-based encoder, they are directed to four distinct proxy heads, each responsible for a specific pretraining task. The proxies include one symmetry-aware head that identifies the symmetrical features of the input sub-volumes and three heads that deal with intrinsic

features of the input sub-volume (i.e., inpainting, rotation, and contrastive coding) as shown in Figure 2.

The Symmetry-Aware Head (SAH) consists of a max pooling layer followed by a multi-layer perceptron (MLP), which distills the features for an original input patch and the symmetrical counterpart as $\mathbf{f}^s$ and $\tilde{\mathbf{f}}^s$. These features serve as the foundation of computing the conditional symmetry loss, expressed as:

$$\mathcal{L}_{Symmetry} = -\log \frac{\sum \exp(\text{sim}(\mathbf{f}^s, \tilde{\mathbf{f}}^s)/\tau) \cdot y^0}{\sum \exp(\text{sim}(\mathbf{f}^s, \tilde{\mathbf{f}}^s)/\tau)}, \tag{2}$$

where y^0 indicates the subject's health status, with 1 for healthy control and 0 for patients. The parameter τ is a temperature factor adjusting the logits' scaling. The cosine similarity $\text{sim}(\mathbf{a}, \mathbf{b})$ is used to measure the similarity between embedded features. The core premise of this formulation is that for subjects with healthy brains, a given patch and its symmetrical counterpart should exhibit similar visual and structural characteristics, whereas deviations and asymmetries are expected in the presence of pathological conditions, reflecting the brain's asymmetrical response to disease.

The heads for intrinsic feature awareness follow [27], including the Inpainting Head (IH), defined as a decoder to reconstruct the occluded voxels; the Rotation Head (RH), defined by an MLP classifier to identify the rotation angles; and the Contrastive Coding Head (CH), defined by two MLP layers to distinguish the present sub-volumes from other sub-volumes. The losses are defined as:

$$\begin{aligned} &\mathcal{L}_{Inpainting} = \|\mathbf{f}^p - \mathbf{x}\|_1 \\ &\mathcal{L}_{Rotation} = -r \log(\text{Softmax}(\mathbf{f}^r)), \\ &\mathcal{L}_{Contrastive} = -\log \frac{\exp(\text{sim}(\mathbf{f}_i^c, \mathbf{f}_j^c)/\tau)}{\sum \mathbb{1}_{k \neq i} \exp(\text{sim}(\mathbf{f}_i^c, \mathbf{f}_k^c)/\tau)}, \end{aligned} \tag{3}$$

where $\mathbf{f}^p, \mathbf{f}^r, \mathbf{f}^c$ represent the concatenated output from each head, and $\mathbf{x}$ is the input sub-volume. Through comprehensive grid search experimentation, the four losses are configured to be evenly summed up for pretraining, balancing the contribution of each task effectively. The pretrained symmetry-aware network can be further applied for downstream tasks, as shown in Section 3.

3 Experiments and Results

We evaluated our proposed approach by pretraining on large neuroimaging datasets and downstream classification and segmentation tasks.

3.1 Datasets and Downstream Tasks

We curated two large datasets for brain image pretraining using head CT and brain MRI images separately. The MRI dataset comprises 3D T1 images of 3,509 healthy adults and 2,917 patients sourced from a variety of publicly available datasets [4,29,24,23,19,16,3,16] (more details in supplementary). The CT

dataset is derived from our proprietary in-house collection, including 2025 images from healthy individuals and 4693 images from patients afflicted with a range of brain diseases, with images from three brain diseases (i.e., hemorrhage, fracture, midline shift) reserved for downstream tasks.

For MRI-based evaluations, we assess the zero-shot and few-shot classification capabilities of our model across several disease identification tasks, including Focal Epilepsy [25], ADHD [5], and Schizophrenia [26]. Furthermore, we explore the model's transfer learning proficiency in segmentation tasks using images with different modalities from the MRI pretraining. This includes the BraTS2021 dataset [3] and the MSSEG2016 dataset [8].

For CT imaging, we conducted few-shot training using reserved in-house data of three brain diseases, including intracranial hemorrhage, bone fracture, and midline shift, and evaluated the performance on the public dataset CQ500 [7] to demonstrate the model's versatility across varied clinical conditions.

3.2 Implementation Details

The proposed model undergoes pretraining with distinct settings for image classification and segmentation tasks. Specifically, the whole image classification task pretraining is conducted with a batch size of 2 and images resized to $160 \times 160 \times 160$, and a batch size of 4 and sub-volume dimensions of $96 \times 96 \times 96$ was used for segmentation-related pretraining. The pretraining phase is executed with a learning rate of 6×10^{-6} over 300 epochs. For downstream tasks, the learning rate is set to 1×10^{-5}, spanning 100 epochs. All models are trained utilizing NVIDIA Tesla V100 GPUs.

The pretrained models are further experimented with downstream tasks under a 5-fold cross-validation approach across all experiments, with the exception of the multiple sclerosis lesion segmentation on MSSEG, which has a predefined validation set [21]. The average performance scores, including Area under the ROC Curve (AUC), F1-measure, and Dice Similarity Coefficient (DSC), are reported.

3.3 Results

The classification outcomes are delineated in Tables 1 and 2. In the supervised fine-tuning scenario (checked under "+SFT"), our model consistently surpassed the performance of the Swin MLP [27] across all evaluated downstream tasks, especially in some diseases that display significant asymmetrical features, such as ADHD and Midline Shift. Besides, our method has achieved comparable performance in CT disease diagnosis as that of [6], but only hundreds of labeled cases were used for fine-tuning, much less than in [6] where 313,318 scans were collected and manually labeled. This underscores the significant advantage of incorporating symmetrical awareness during the pretraining phase.

Furthermore, we assessed the impact of symmetrical awareness in a zero-shot learning context. Remarkably, our pretrained model exhibited superior zero-shot

Table 1: Benchmark methods on the MRI T1-only classification tasks.

Experiments	+Pretrain	+SACA	+SFT	Focal Epilepsy AUC	Focal Epilepsy F1	Schizophrenia AUC	Schizophrenia F1	ADHD AUC	ADHD F1
Swin MLP [27]	✗	✗	✓	0.678	0.683	0.656	0.810	0.685	0.625
	✓	✗	✗	0.522	0.651	0.594	0.749	0.514	0.609
	✓	✗	✓	0.713	0.695	0.627	0.803	0.642	0.645
Ours	✓	✓	✗	0.631	0.683	0.663	0.760	0.515	0.625
	✓	✓	✓	**0.778**	**0.788**	**0.719**	**0.836**	**0.792**	**0.727**

capabilities in almost all tasks. This indicates that our pretraining strategy effectively captures critical pathological features based on symmetry characteristics, facilitating generalization to novel tasks without requiring task-specific data.

Table 2: Benchmark methods on the CT classification tasks evaluated on the CQ500 dataset [6].

Experiments	+Pretrain	+SACA	+SFT	Hemorrhage AUC	Hemorrhage F1	Fracture AUC	Fracture F1	Midline Shift AUC	Midline Shift F1
Qure.ai [6]	✗	✗	✓	0.942	0.761	0.962	0.508	0.970	0.704
Swin MLP [27]	✗	✗	✓	0.872	0.776	0.918	0.460	0.926	0.640
	✓	✗	✗	0.751	0.685	0.781	0.353	0.788	0.413
	✓	✗	✓	0.911	0.748	0.952	0.469	0.942	0.620
Ours	✓	✓	✗	0.793	0.722	0.803	0.267	0.866	0.508
	✓	✓	✓	**0.958**	**0.781**	**0.969**	**0.513**	**0.986**	**0.728**

The assessment of models' proficiency in adapting to segmentation tasks is summarized in Table 3, which manifests the superior performance of our method. Additionally, we conducted an ablation study to discern the relative contribution of our proposed SACA module and SAH proxy. Both the SAH only pretraining process and the SACA module could improve the performance of segmentation steadily, and the combination of all pretraining proxies and the novel SACA module could further yield the most substantial benefits, underscoring the synergetic effect of combining these diverse learning strategies. This comprehensive evaluation demonstrates our method's superior capability in leveraging pretraining proxies to enhance model performance on segmentation tasks, highlighting its potential for advancing medical image analysis.

4 Conclusion

Our study introduces an innovative framework utilizing 3D brain imaging data of various modalities, significantly enhancing neuroimage analysis by encoding symmetry-aware features in deep neural networks. The integration of the

Table 3: Benchmark methods and ablation study of our method on the MRI segmentation tasks. "IRC" represents commonly used pretrain heads for inpainting (I), rotation (R), and contrastive (C).

Experiments				BraTS [3]				MSSEG [8]
	+IRC	+SAH	+SACA	Dice_{ET}	Dice_{TC}	Dice_{WT}	Dice_{Avg}	Dice
nnU-Net [15]	✗	✗	✗	0.883	0.927	0.913	0.908	-
TransBTS [28]	✗	✗	✗	0.868	0.911	0.898	0.891	-
SwinUNETR [11]	✗	✗	✗	0.891	0.933	0.917	0.913	0.594
Ours	✗	✓	✗	0.897	0.935	0.920	0.917	0.615
	✗	✗	✓	0.901	0.937	0.928	0.922	0.640
	✗	✓	✓	0.909	0.940	0.935	0.928	0.654
	✓	✓	✓	**0.912**	**0.946**	**0.940**	**0.932**	**0.680**

Symmetry-Aware Cross-Attention (SACA) module and the symmetry-aware pretraining proxy has shown remarkable improvements in classification and segmentation tasks, setting new benchmarks over existing models. This work highlights the critical role of anatomical symmetry in neuroimaging analysis and suggests a promising direction for developing more accurate, robust, and generalized AI models in healthcare leveraging intrinsic anatomical features.

Acknowledgments. The authors would like to express their gratitude to the BCR Scholarship for Computational Neuroscience and the Faculty of Engineering Research Stipend Scholarship at the University of Sydney for their financial support, which enabled this research and its contributions to medical imaging and healthcare.

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

References

1. Atito, S., Awais, M., Kittler, J.: Sit: Self-supervised vision transformer. arXiv preprint arXiv:2104.03602 (2021)
2. Azizi, S., Mustafa, B., Ryan, F., Beaver, Z., Freyberg, J., Deaton, J., Loh, A., Karthikesalingam, A., Kornblith, S., Chen, T., et al.: Big self-supervised models advance medical image classification. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 3478–3488 (2021)
3. Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., et al.: The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv preprint arXiv:2107.02314 (2021)
4. Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., et al.: Toward discovery science of human brain function. Proceedings of the National Academy of Sciences **107**(10), 4734–4739 (2010)
5. Booth, J.R., Cooke, G., et al.: Working memory and reward in children with and without attention deficit hyperactivity disorder (adhd) (2021). `https://doi.org/10.18112/openneuro.ds002424.v1.2.0`

6. Chilamkurthy, S., Ghosh, R., Tanamala, S., Biviji, M., Campeau, N.G., Venugopal, V.K., Mahajan, V., Rao, P., Warier, P.: Development and validation of deep learning algorithms for detection of critical findings in head ct scans. arXiv preprint arXiv:1803.05854 (2018)
7. Chilamkurthy, S., Ghosh, R., Tanamala, S., Biviji, M., et al.: Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study. The Lancet **392**(10162), 2388–2396 (2018)
8. Commowick, O., Kain, M., Casey, R., Ameli, R., Ferré, J.C., Kerbrat, A., Tourdias, T., Cervenansky, F., Camarasu-Pop, S., Glatard, T., et al.: Multiple sclerosis lesions segmentation from multiple experts: The miccai 2016 challenge dataset. Neuroimage **244**, 118589 (2021)
9. Gidaris, S., Singh, P., Komodakis, N.: Unsupervised representation learning by predicting image rotations. arXiv preprint arXiv:1803.07728 (2018)
10. Haghighi, F., Taher, M.R.H., Zhou, Z., Gotway, M.B., Liang, J.: Transferable visual words: Exploiting the semantics of anatomical patterns for self-supervised learning. IEEE transactions on medical imaging **40**(10), 2857–2868 (2021)
11. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: International MICCAI Brainlesion Workshop. pp. 272–284. Springer (2021)
12. Herzog, N.J., Magoulas, G.D.: Deep learning of brain asymmetry images and transfer learning for early diagnosis of dementia. In: International Conference on Engineering Applications of Neural Networks. pp. 57–70. Springer (2021)
13. Hua, Y., Yan, Z., Kuang, Z., Zhang, H., Deng, X., Yu, L.: Symmetry-aware deep learning for cerebral ventricle segmentation with intra-ventricular hemorrhage. IEEE Journal of Biomedical and Health Informatics **26**(10), 5165–5176 (2022)
14. Illán, I.A., Górriz, J.M., Ramírez, J., et al.: Bilateral symmetry aspects in computer-aided alzheimer's disease diagnosis by single-photon emission-computed tomography imaging. Artificial intelligence in medicine **56**(3), 191–198 (2012)
15. Isensee, F., Jäger, P.F., Full, P.M., Vollmuth, P., Maier-Hein, K.H.: nnu-net for brain tumor segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6. pp. 118–132. Springer (2021)
16. Jack Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., et al.: The alzheimer's disease neuroimaging initiative (adni): Mri methods. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **27**(4), 685–691 (2008)
17. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage **17**(2), 825–841 (2002)
18. Kuo, F., Massoud, T.F.: Structural asymmetries in normal brain anatomy: A brief overview. Annals of Anatomy-Anatomischer Anzeiger **241**, 151894 (2022)
19. LaMontagne, P.J., Benzinger, T.L., Morris, J.C., Keefe, S., Hornbeck, R., Xiong, C., Grant, E., Hassenstab, J., Moulder, K., Vlassenko, A.G., et al.: Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. MedRxiv pp. 2019–12 (2019)
20. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., et al.: Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 10012–10022 (2021)

21. Ma, Y., Zhang, C., Cabezas, M., Song, Y., Tang, Z., Liu, D., Cai, W., Barnett, M., Wang, C.: Multiple sclerosis lesion analysis in brain magnetic resonance images: techniques and clinical applications. IEEE Journal of Biomedical and Health Informatics **26**(6), 2680–2692 (2022)
22. Narr, K.L., Bilder, R.M., Luders, E., Thompson, P.M., Woods, R.P., Robinson, D., Szeszko, P.R., et al.: Asymmetries of cortical shape: effects of handedness, sex and schizophrenia. Neuroimage **34**(3), 939–948 (2007)
23. Nooner, K.B., Colcombe, S.J., Tobe, R.H., Mennes, M., Benedict, M.M., Moreno, A.L., Panek, L.J., Brown, S., Zavitz, S.T., Li, Q., et al.: The nki-rockland sample: a model for accelerating the pace of discovery science in psychiatry. Frontiers in neuroscience **6**, 152 (2012)
24. Rowland, A., Burns, M., Hartkens, T., Hajnal, J., Rueckert, D., Hill, D.L.: Information extraction from images (ixi): Image processing workflows using a grid enabled image database. Proceedings of DiDaMIC **4**, 55–64 (2004)
25. Schuch, F., Walger, L., Schmitz, M., et al.: An open presurgery mri dataset of people with epilepsy and focal cortical dysplasia type ii (2023). `https://doi.org/10.18112/openneuro.ds004199.v1.0.5`
26. Soler-Vidal, J., Fuentes-Claramonte, P., Salgado-Pineda, P., Ramiro, N., García-León, M.Á., Torres, M.L., Arévalo, A., Guerrero-Pedraza, A., Munuera, J., Sarró, S., et al.: Brain correlates of speech perception in schizophrenia patients with and without auditory hallucinations. PloS one **17**(12), e0276975 (2022)
27. Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V., Hatamizadeh, A.: Self-supervised pre-training of swin transformers for 3d medical image analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20730–20740 (2022)
28. Wang, W., Chen, C., et al.: Transbts: Multimodal brain tumor segmentation using transformer. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24. pp. 109–119. Springer (2021)
29. Zuo, X.N., Anderson, J.S., Bellec, et al.: An open science resource for establishing reliability and reproducibility in functional connectomics. Scientific Data **1**(1), 1–13 (2014)