HyDA: Hypernetworks for Test Time Domain Adaptation in Medical Imaging Analysis

Doron Serebro[1[0009−0000−3941−5017]] and Tammy Riklin-Raviv[1[0000−0003−2532−5875]]

The School of Electrical and Computer Engineering and the Data Science Research Center, Ben-Gurion University of The Negev
doronser@post.bgu.ac.il,rrtammy@ee.bgu.ac.il

Abstract. Medical imaging datasets often vary due to differences in acquisition protocols, patient demographics, and imaging devices. These variations in data distribution, known as domain shift, present a significant challenge in adapting imaging analysis models for practical healthcare applications. Most current domain adaptation (DA) approaches aim either to align the distributions between the source and target domains or to learn an invariant feature space that generalizes well across all domains. However, both strategies require access to a sufficient number of examples, though not necessarily annotated, from the test domain during training. This limitation hinders the widespread deployment of models in clinical settings, where target domain data may only be accessible in real time.

In this work, we introduce HyDA, a novel hypernetwork framework that leverages domain-specific characteristics rather than suppressing them, enabling dynamic adaptation at inference time. Specifically, HyDA learns implicit domain representations and uses them to adjust model parameters on-the-fly, allowing effective interpolation to unseen domains. We validate HyDA on two clinically relevant applications—MRI-based brain age prediction and chest X-ray pathology classification—demonstrating its ability to generalize across tasks and imaging modalities. Our code is available at: https://github.com/doronser/hyda.

Keywords: Domain Adaptation · Hypernetworks · MRI · X-Ray

1 Introduction

Deep learning has significantly advanced medical image analysis, enabling accurate detection, classification, segmentation, and predictive modeling. However, for practical deployment in healthcare, models must adapt to variations in imaging protocols, scanner types, and patient demographics, which lead to discrepancies between training and test data distributions. This issue, known as domain shift, remains a major barrier to robust and generalizable model performance [10].

Current domain adaptation (DA) techniques generally aim to either align the distributions between the source and target domains [34,31] or learn a consistent feature space across different domains [20,35,9]. Yet, both approaches depend on having enough target domain samples during training, whether annotated or not. This dependency poses a challenge for the deployment of models in clinical settings, where the target data may only be available at the time of the test. In this work, we propose HyDA, a novel hypernetwork framework that exploits domain characteristics rather than discarding them, enabling dynamic adaptation during both training and inference. Specifically, HyDA learns implicit domain representations that are used to generate weights and biases for a primary network on-the-fly, effectively interpolating to unseen domains. HyDa is task and modality-agnostic, making it easily integrable into various medical imaging applications. We showcase its generality and robustness for two clinically relevant tasks—chest X-ray pathology classification and MRI brain age prediction, demonstrating superior performance over baseline and other domain adaptation techniques.

2 Related Works

Domain Adaptation Methods. Unsupervised domain adaptation (UDA) addresses shifting data distributions between source and target domains. One prominent approach is domain adversarial learning, as exemplified by Domain-Adversarial Neural Networks (DANN) [8], while MDAN (Multi-Domain Adversarial Network) extends this idea to multiple source domains [38]. Another line of work focuses on invariant feature learning; for example, Deep CORrelation ALignment (CORAL) minimizes domain discrepancy by aligning the second-order statistics of source and target features [30].
Recently, transformer-based methods have gained traction for their self-attention capabilities. TransDA leverages domain-specific tokens and cross-attention to align features in an unsupervised manner [36]. Similarly, AdaptFormer [3] and DAFormer [13] integrate lightweight adapter modules within a Vision Transformer framework to modulate representations based on domain cues while maintaining a shared global representation.

Test-time domain adaptation (TTDA) techniques have a key advantage over the methods mentioned above: they do not require target data during training and can adapt models on-the-fly during inference. For example, TENT (Test Time Entropy Minimization) adjusts model parameters via entropy minimization, which works well for multi-class classification tasks with clear output probabilities [31]. MEMO stabilizes adaptation under distribution shifts using augmentations [37]. Although not strictly a TTDA method, SHOT (Source Hypothesis Transfer) adapts to target data without requiring source samples [17] by relying on pseudo-labeling and entropy minimization. However, the reliance on entropy may limit their applicability to tasks such as regression or multi-label classification without modifications.

Hypernetworks. First introduced by Ha et al. [11], hypernetworks are neural networks that generate weights and biases for primary networks, dynamically creating a unique set of parameters for each input. Their effectiveness has been demonstrated in various tasks, including 3D shape reconstruction [18], federated

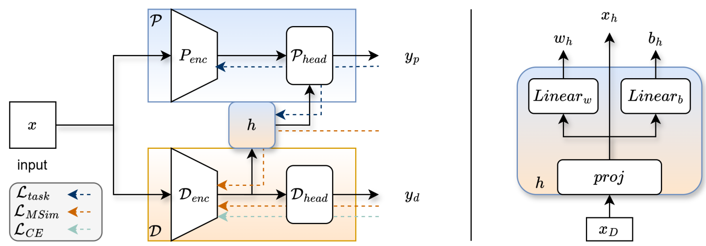


Fig. 1: The proposed HyDA framework (left) is composed of a hypernetwork h (right), a primary network $\mathcal{P}$, and a domain classifier $\mathcal{D}$. The hypernetwork generates weights and biases to the primary's head $\mathcal{P}_{head}$ based of the domain feature vector X_D provided by the domain encoder $\mathcal{D}_{enc}$. Other weights in the system are internal and are learned via back-propagation using a regularized task dependent $\mathcal{L}_{RT}$, classification $\mathcal{L}_{CE}$ or/and multi-similarity $\mathcal{L}_{MSim}$ loss functions as illustrated by the dashed arrows.

learning [29], and medical image segmentation [21]. Aharon et al. [1] showed that hypernetworks can interpolate by conditioning an image denoising model on expected noise variance, while Duenias et al. [6] used them to condition medical image analysis on tabular data. Building on these ideas, we show that hypernetworks can be applied to medical imaging domain adaptation by generating weights from domain features, effectively interpolating across the domain space.

3 Method

Our proposed HyDA framework, illustrated in Fig. 1, is composed of a primary network $\mathcal{P}$ that could have any architecture addressing any medical imaging analysis task; a hypernetwork h and a domain classifier $\mathcal{D}$. Being trained on datasets from different source domains - the classifier learns implicit domain features that are mapped by h to sets of weights and biases. These parameters, termed *external parameters* are transferred to a subset of layers in $\mathcal{P}$.

Formally, let $x \in \mathbb{R}^d$, denote a $d-$dimensional input image ($d \in \{2, 3\}$). We define by $\mathcal{D}_{enc}$ and $\mathcal{D}_{head}$ the domain encoder and domain head, respectively, which together compose the classifier $\mathcal{D}$.

The classifier is trained to predict the domain of a training image x as follows:

$$y_D = \mathcal{D}_{head}(\mathcal{D}_{enc}(x)) \tag{1}$$

where y_D is the label of a source domain. We assume the availability of at least two source domains. If deployed separately, the trained domain encoder maps any input x into a domain feature vector, i.e.,

$$x_D = \mathcal{D}_{enc}(x) \tag{2}$$

where $x_D \in f_D$ is the domain feature vector of x and f_D denotes the domain feature space. The primary network $\mathcal{P}$ which is trained to predict an output y_P for x can be formalized as follows:

$$y_P = \mathcal{P}_{head}(\mathcal{P}_{enc}(x), h(D_{enc}(x))), \tag{3}$$

where, $\mathcal{P}_{enc}$ denotes the *internal* layers in $\mathcal{P}$ which are trained through a standard back propagation process and $h(D_{enc}(x))$ defines its *external* domain-aware weights and biases - generated by the hypernetwork h.

We hypothesize that during inference, feature vectors of a target domain x_D, unseen by the domain encoder before, are well embedded within the domain feature space, f_D, and can be represented as linear combinations of training domain features. We aim to optimize the hypernetwork such that the metric capturing inter-domain and intra-domain relationships within f_D is preserved in the external primary network parameters. Once optimally converged, HyDA can interpolate to new target domains at test time.

3.1 Domain Conditioning Hypernetwork

The hypernetwork maps the domain embedding x_D to weights and biases (w_h, b_h) for the external primary network layers. Let N, O, B denote the layer's input, output and batch size, respectively. In a standard linear layer, the output $\psi \in \mathbb{R}^{(B,O)}$ is computed as:

$$\psi = \chi * w, \quad w \in \mathbb{R}^{(N,O)}$$

where $\chi \in \mathbb{R}^{(B,N)}$ is the input batch and $*$ is matrix multiplication. A hyper-linear layer instead assigns a unique weight matrix to each batch element:

$$\psi_i = \chi_i * w_h^i, \quad w_h^i \in \mathbb{R}^{(N,O)} \quad i = 1, \dots, B$$

The hypernetwork is flexible and can be implemented in various ways. For simplicity, we use a single linear layer followed by a ReLU activation, which is sufficient to generate effective domain-aware weights for the primary network.

To ensure stable convergence, we initialize the hypernetwork weights as in Chang et al. [2] such that the input variance is preserved in the primary network. We also regularize the weights using the l_2 norm.

3.2 Loss Functions

The hypernetwork and the internal primary network layers are trained in an end-to-end manner with a regularized loss function as follows:

$$\mathcal{L}_{RT} = \mathcal{L}_{task} + \lambda_{BP} \left\| w_{BP} \right\|_2 + \lambda_h \left\| w_h \right\|_2 \tag{4}$$

where, RT stands for regularized task, BP for backpropagation, $\mathcal{L}_{task}$ is a task-dependent loss (e.g. cross-entropy for classification, MSE for regression), w_{BP} denote the union of the hypernetwork's and the internal primary network's weights,

w_h are the external primary network weights, generated by the hypernetwork, and λ_{BP}, λ_h are their corresponding coefficients. The domain classifier is trained using the following loss:

$$\mathcal{L}_D = \mathcal{L}_{CE} + \alpha \mathcal{L}_{MSim} + \lambda_D \left\| w_D \right\|_2 \tag{5}$$

where $\mathcal{L}_{CE}$ is the cross-entropy loss, $\mathcal{L}_{MSim}$ is multi-similarity loss as in Wang et. al. [33], w_D are the domain network's weights and α, λ_D are coefficients. The multi-similarity loss optimizes over hard positive and negative examples. As domain labels are known during training, these samples can be selected directly. The supervised CE loss $\mathcal{L}_{CE}$ aims to correctly classify the input into source domains, while the contrastive, multi-similarity loss encourages the separation of embedded domain feature vectors into different domain-aware clusters. The multi-similarity loss also supports the hypernetwork training - allowing it to maintain domain-specific representation of the weights and biases it generates for the primary network.

4 Experiments and Results

We demonstrate the proposed HyDA framework on two medical imaging analysis tasks: chest X-ray pathology classification and MRI brain age prediction. In both experiments, we use a leave-one-out setting to simulate different source and target domain configurations. Specifically, given N datasets, training is performed on $N-1$ domains and testing on the held-out domain, cycling through all domains.

4.1 Chest X-ray Pathology Classification

We trained our model for multi-label classification on chest X-ray scans from three publicly available datasets, comparing HyDA to a baseline with no adaptation, a UDA method (soft MDAN [38]), and a TTDA method (TENT [31]). In both cases, the domain classifier was pre-trained for robust initialization.

Data. We use the NIH [32], CheXpert [14], and VinDr [24] datasets, selecting five classes—Atelectasis, Cardiomegaly, Consolidation, Effusion, and Pneumothorax—that are common across all three, resulting in a combined dataset of $90,570$ X-ray scans.

Implementation Details. We fine-tuned a DenseNet121 model pre-trained on ImageNet, replacing its input and output layers to process single-channel images and output five classes, following prior work [27,4]. The domain classifier is a simple CNN with four convolution blocks and a linear classification layer, while the hypernetwork is a multi layer perceptron (MLP) that generates a set of weights and biases for the primary network (DenseNet). Both baseline and HyDA models were trained for 150 epochs using the AdamW optimizer (learning rate: $1e-3$, weight decay: 0.05) with a cosine annealing scheduler (minimum learning rate: $1e-6$).

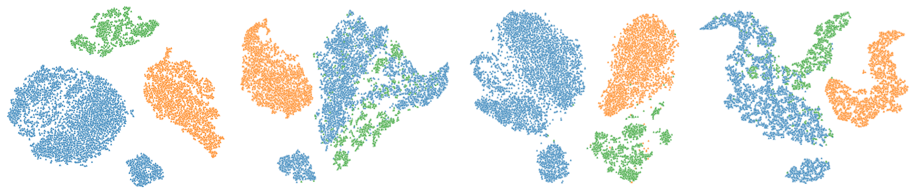
Target Domain	Method	Pathologies (AUC) ↑					Avg. (std)
		Atel.	Cardio.	Cons.	Eff.	Pneu.	
-	Baseline	0.85	0.95	0.86	0.94	0.87	0.89 (0.04)
	MDAN	0.86	0.96	0.86	0.94	0.88	0.90 (0.04)
	HyDA	**0.87**	**0.97**	**0.86**	**0.94**	**0.89**	**0.91 (0.04)**
NIH	Baseline	**0.70**	0.81	**0.76**	0.86	0.77	0.78 (0.06)
	MDAN	0.67	0.89	0.76	0.86	0.77	0.79 (0.08)
	TENT	0.61	0.70	0.64	0.81	0.67	0.69 (0.07)
	HyDA	0.68	**0.89**	0.75	**0.88**	**0.79**	**0.80 (0.08)**
CheXpert	Baseline	0.81	**0.86**	0.73	0.87	0.74	0.80 (0.06)
	MDAN	0.77	0.76	0.71	0.84	0.72	0.76 (0.05)
	TENT	0.76	0.86	0.77	0.89	0.76	0.81 (0.06)
	HyDA	**0.82**	0.85	**0.82**	**0.89**	**0.74**	**0.82 (0.05)**
VinDr	Baseline	0.60	0.76	0.85	0.88	0.91	0.80 (0.11)
	MDAN	**0.68**	0.82	0.88	0.87	0.89	0.83 (0.08)
	TENT	0.51	0.72	0.80	0.74	0.86	0.73 (0.12)
	HyDA	0.66	**0.87**	**0.93**	**0.89**	**0.92**	**0.85 (0.10)**

Table 1: Chest X-ray classification results measured by AUC. Pathologies abbreviations: Atel (Atelectasis), Cardio (Cardiomegaly), Cons (Consolidation), Eff (Effusion), Pneu (Pneumothorax). Each group compares different models on the same target domain. Best results in bold.

Results. Table 1 reports the area under curve (AUC) of the chest X-ray experiments. HyDA outperforms the baseline in both fully supervised and leave-one-out settings. Notably, the improvement correlates with the separability of domain features that were not seen in training (see Fig. 2); domains with well-clustered features (CheXpert and VinDr) show larger gains compared to NIH. Paired t-tests comparing HyDA to the comparison methods show the results are statistically significant - p-values are 0.0023, 0.0010, and 0.00007 when comparing to baseline, MDAN and TENT respectively — all p-values are below 0.0025.
Ablation Study. Table 2 highlights the contribution of each component in the proposed loss function to achieve the best possible performance.

4.2 Brain Age Prediction

To further assess our method and demonstrate its task-agnostic nature, we evaluated its performance on age prediction from brain MRI scans.
Data. We used 18 brain MRI datasets containing 26,691 scans. The scans were preprocessed using the workflow in Levakov et. al. [16].
Implementation Details. Our primary network is a 3D CNN comprised of 4 convolution blocks followed by a 4-layered-MLP. The domain classifier follows a similar architecture, with fewer parameters (refer to our code for further details). The model was trained using AdamW optimizer with a learning rate of $1e-$



(a) All domains (b) w/o NIH (c) w/o CheXpert (d) w/o VinDr

Fig. 2: t-SNE projections of domain feature in fully supervised and leave-one-out settings. The plots illustrate the embedding of previously unseen domains in the learned domain feature space : (a) All domains training with respect to training (b) w/o NIH (blue) (c) w/o CheXpert (orange) and (d) w/o VinDr (green).

$\mathcal{L}_{CE}$	$\mathcal{L}^D_{MSim}$	$\mathcal{L}^h_{MSim}$	**NIH**	**CheXpert**	**VinDr**
✓			0.72 (0.11)	0.79 (0.05)	0.81 (0.13)
✓	✓		0.76 (0.09)	0.81 (0.06)	0.83 (0.10)
✓	✓	✓	**0.80 (0.08)**	**0.82 (0.05)**	**0.85 (0.10)**

Table 2: Ablation study of the loss terms. Each row represents an incremental combination of loss terms, including domain classifier's cross-entropy (CE) $\mathcal{L}_{CE}$ and multi-similarity (MSim) $\mathcal{L}^D_{MSim}$ loss functions as well as hypernetworks' MSim loss $\mathcal{L}^h_{MSim}$. Average AUC results (std in brackets) of the target domain for each of the three datasets are reported.

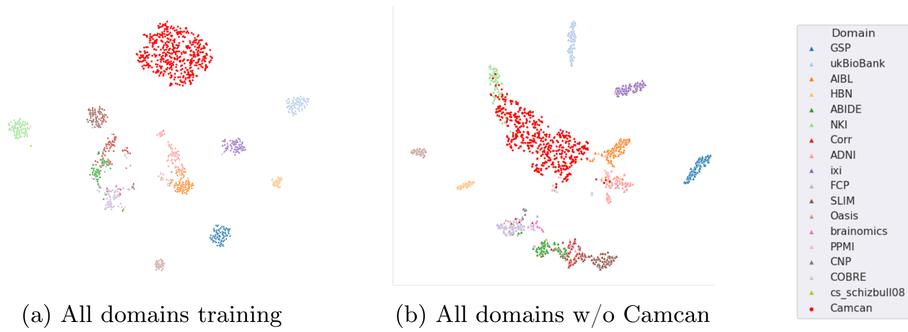
4, weight decay of 0.05 and a cosine annealing learning rate scheduler with a minimum learning rate of $1e-6$ for 150 epochs.

Results. Table 3 presents the brain age prediction results. Notably, HyDA outperforms the baseline in both supervised and leave-one-out settings. These results demonstrate HyDA's ability to learn meaningful domain representations, interpolate to unseen domains, and adapt the model on-the-fly using domain features. The interpolation capability is further illustrated in the t-SNE plot in Fig. 3, where samples from a previously unseen domain are well embedded among feature vectors from domains used during training. Paired t-tests comparing HyDA with the baseline show that, while the supervised setting results are not statistically significant, the leave-one-out results are significant (p-values of 0.1122 and 0.0383, respectively).

Ablation Study. We evaluated the robustness of HyDA by testing different configurations of the primary network's MLP head, which consists of four layers—three of which can be external, with their weights generated by the hypernetwork. Table 4 shows that replacing some internal layers with domain-specific (external) ones improves performance over the baseline, regardless of which layers are adapted. The best performance is achieved by combining both types of layers: relying solely on task-specific weights limits generalization to unseen do-

Model	CNP [26]	NKI [25]	ixi [12]	Oasis [22]	ABIDE [5]	ADNI [15]	AIBL [7]	PPMI [23]	Camcan [28]	SLIM [19]	Avg. (std)
Fully Supervised (Validation MAE) ↓											
Baseline	3.11	3.01	3.54	**3.29**	2.09	**2.80**	**2.74**	4.23	3.35	0.47	2.86 (0.96)
HyDA	**2.39**	**2.92**	**3.22**	**3.29**	**1.74**	3.04	2.94	**3.94**	**3.21**	**0.37**	**2.71 (0.95)**
Leave-One-Out (Test MAE) ↓											
Baseline	3.36	3.90	4.41	5.40	3.25	**4.31**	3.56	**4.15**	3.50	1.44	3.73 (0.97)
HyDA	**2.86**	**3.44**	**4.14**	**5.20**	**3.16**	4.48	**3.45**	4.24	**3.35**	**1.34**	**3.57 (1.00)**

Table 3: Brain age prediction results in fully supervised (validation MAE) and leave-one-out (test MAE) settings. Best scores are in bold.



(a) All domains training (b) All domains w/o Camcan

Fig. 3: t-SNE projections of domain feature vectors x_D, Camcan examples are in red. The plots show embedding when samples of all domains (a) or all but Camcan (b) are available during training.

Layer 1	Layer 2	Layer 3	Average (std)
			4.16 (0.26)
✓			3.99 (0.20)
	✓		3.97 (0.32)
		✓	3.79 (0.21)
	✓	✓	3.79 (0.35)
✓	✓	✓	4.17 (0.11)

Table 4: Hypernetwork external layer configuration - ablation study. Each configuration was trained on two target domains (NKI, ixi), and results are reported as mean (std) target domain MAE.

mains, while using only domain-specific weights compromises critical task-related information.

5 Conclusions

We introduced HyDA, a hypernetwork-based framework that rethinks test-time domain adaptation in medical imaging by embracing domain variability rather than suppressing it. By learning implicit domain representations and dynamically generating model parameters at test time, HyDA tailors predictions for each input based on its domain characteristics.

Experimental evaluations on chest X-ray pathology classification and MRI brain age prediction demonstrate that HyDA outperforms traditional domain-invariant methods and existing test-time adaptation techniques. Its ability to interpolate between domains, as revealed by t-SNE visualizations, confirms that leveraging domain-specific cues leads to more robust and generalizable models. Moreover, HyDA's task-agnostic design and compatibility with various architectures make it a versatile solution for a wide range of clinical applications. Overall, HyDA offers a promising pathway toward more reliable and adaptable medical image analysis, enabling models to seamlessly adjust to real-world variations in data acquisition without requiring extensive target domain training.

Acknowledgments

This work was partially supported by the Israel Ministry of Science and Technology (T.R.R. MOST 0007311/1001817977).

References

1. Aharon, S., Ben-Artzi, G.: Hypernetwork-based adaptive image restoration. In: International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1–5 (2023)
2. Chang, O., Flokas, L., Lipson, H.: Principled weight initialization for hypernetworks. In: International Conference on Learning Representations (2019)
3. Chen, S., Ge, C., Tong, Z., et al.: Adaptformer: Adapting vision transformers for scalable visual recognition. Advances in Neural Information Processing Systems (NeurIPS) **35**, 16664–16678 (2022)
4. Cohen, J.P., Hashir, M., Brooks, R., Bertrand, H.: On the limits of cross-domain generalization in automated x-ray prediction. In: Medical Imaging with Deep Learning. pp. 136–155. PMLR (2020)
5. Di Martino, A., Yan, C.G., Li, Q., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular psychiatry **19**(6), 659–667 (2014)
6. Duenias, D., Nichyporuk, B., Arbel, T., Riklin Raviv, T.: Hyperfusion: A hypernetwork approach to multimodal integration of tabular and medical imaging data for predictive modeling. arXiv preprint arXiv:2403.13319 (2024)
7. Ellis, K.A., Bush, A.I., Darby, D., et al.: The australian imaging, biomarkers and lifestyle (aibl) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of alzheimer's disease. International psychogeriatrics **21**(4), 672–687 (2009)
8. Ganin, Y., Ustinova, E., Ajakan, H., et al.: Domain-adversarial training of neural networks. Journal of machine learning research **17**(59), 1–35 (2016)
9. Gu, M., Thies, M., Mei, S., et al.: Unsupervised domain adaptation using soft-labeled contrastive learning with reversed monte carlo method for cardiac image segmentation. In: Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 681–691. Springer (2024)

10. Guan, H., Liu, M.: Domain adaptation for medical image analysis: a survey. IEEE Transactions on Biomedical Engineering **69**(3), 1173–1185 (2021)
11. Ha, D., Dai, A.M., Le, Q.V.: Hypernetworks. In: International Conference on Learning Representations (2017)
12. Heckemann, R.A., Hartkens, T., Leung, K.K., et al.: Information extraction from medical images: developing an e-science application based on the globus toolkit. In: UK e-Science All Hands Meeting (2003)
13. Hoyer, L., Dai, D., Van Gool, L.: Daformer: Improving network architectures and training strategies for domain-adaptive semantic segmentation. In: Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9924–9935 (2022)
14. Irvin, J., Rajpurkar, P., Ko, M., et al.: Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In: AAAI conference on artificial intelligence. vol. 33, pp. 590–597 (2019)
15. Jack Jr, C.R., Bernstein, M.A., Fox, N.C., et al.: The alzheimer's disease neuroimaging initiative (adni): Mri methods. Journal of Magnetic Resonance Imaging **27**(4), 685–691 (2008)
16. Levakov, G., Rosenthal, G., Shelef, I., Riklin Raviv, T., Avidan, G.: From a deep learning model back to the brain—Identifying regional predictors and their relation to aging. Human Brain Mapping **41**(12), 3235–3252 (2020)
17. Liang, J., Hu, D., Feng, J.: Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In: International conference on machine learning. pp. 6028–6039. PMLR (2020)
18. Littwin, G., Wolf, L.: Deep meta functionals for shape representation. In: International Conference on Computer Vision (ICCV). pp. 1824–1833 (2019)
19. Liu, W., Wei, D., Chen, Q., et al.: Longitudinal test-retest neuroimaging data from healthy young adults in southwest china. Scientific data **4**(1), 1–9 (2017)
20. Long, M., Cao, Y., Wang, J., Jordan, M.I.: Learning transferable features with deep adaptation networks. In: International Conference on Machine Learning,ICML. pp. 97–105 (2015)
21. Ma, T., Dalca, A.V., Sabuncu, M.R.: Hyper-convolution networks for biomedical image segmentation. In: Winter Conference on Applications of Computer Vision (WACV). pp. 1933–1942 (2022)
22. Marcus, D.S., Wang, T.H., Parker, J., et al.: Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. Journal of cognitive neuroscience **19**(9), 1498–1507 (2007)
23. Marek, K., Jennings, D., Lasch, S., et al.: The parkinson progression marker initiative (ppmi). Progress in Neurobiology **95**(4), 629–635 (2011), biological Markers for Neurodegenerative Diseases
24. Nguyen, H.Q., Lam, K., Le, L.T., et al.: Vindr-cxr: An open dataset of chest x-rays with radiologist's annotations. Scientific Data **9**(1), 429 (2022)
25. Nooner, K.B., Colcombe, S.J., Tobe, R.H., et al.: The nki-rockland sample: a model for accelerating the pace of discovery science in psychiatry. Frontiers in neuroscience **6**, 152 (2012)
26. Poldrack, R.A., Congdon, E., Triplett, W., et al.: A phenome-wide examination of neural and cognitive function. Scientific data **3**(1), 1–12 (2016)
27. Rajpurkar, P.: Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. ArXiv preprint (2017)
28. Shafto, M.A., Tyler, L.K., Dixon, M., et al.: The cambridge centre for ageing and neuroscience (cam-can) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. BMC neurology **14**, 1–25 (2014)

29. Shamsian, A., Navon, A., Fetaya, E., Chechik, G.: Personalized federated learning using hypernetworks. In: International Conference on Machine Learning (ICML). pp. 9489–9502. PMLR (2021)
30. Sun, B., Saenko, K.: Deep coral: Correlation alignment for deep domain adaptation. In: Computer vision–ECCV 2016 workshops: Amsterdam, part III 14. pp. 443–450. Springer (2016)
31. Wang, D., Shelhamer, E., Liu, S., et al.: Tent: Fully test-time adaptation by entropy minimization. In: International Conference on Learning Representations (2021)
32. Wang, X., Peng, Y., Lu, L., et al.: Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2097–2106 (2017)
33. Wang, X., Han, X., Huang, W., Dong, D., Scott, M.R.: Multi-similarity loss with general pair weighting for deep metric learning. In: Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5022–5030 (2019)
34. Wei, G., Lan, C., Zeng, W., Zhang, Z., Chen, Z.: Toalign: Task-oriented alignment for unsupervised domain adaptation. Advances in Neural Information Processing Systems (NeurIPS) **34**, 13834–13846 (2021)
35. Wolleb, J., Sandkühler, R., Bieder, F., et al.: Learn to Ignore: Domain Adaptation for Multi-site MRI Analysis. In: Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 725–735. Lecture Notes in Computer Science, Springer Nature Switzerland (2022)
36. Yang, G., Tang, H., Zhong, Z., et al.: Transformer-based source-free domain adaptation. arXiv preprint arXiv:2105.14138 (2021)
37. Zhang, M., Levine, S., Finn, C.: Memo: Test time robustness via adaptation and augmentation. Advances in Neural Information Processing Systems (NeurIPS) **35**, 38629–38642 (2022)
38. Zhao, H., Zhang, S., Wu, G., et al.: Adversarial multiple source domain adaptation. Advances in Neural Information Processing Systems (NeurIPS) **31** (2018)