Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 2, April : 2024

DEEP LEARNING BASED HARMONIZATION TECHNIQUES IN MEDICAL IMAGING: A REVIEW

Sankalp Sonu Assistant Professor Department of Computer Science and Engineering Rashtrakavi Ramdhari Singh Dinkar College of Engineering Begusarai, State – Bihar (India), Pin-851134 E-Mail: sankalpsonuiem@gmail.com
Rajiv Kumar Ranjan Assistant Professor Department of Computer Science and Engineering Rashtrakavi Ramdhari Singh Dinkar College of Engineering Begusarai, State – Bihar (India), Pin-851134E-Mail: rajivkr1234@gmail.com
Ankita Sinha Assistant Professor Department of Computer Science and Engineering Rashtrakavi Ramdhari Singh Dinkar College of Engineering Begusarai, State – Bihar (India), Pin-851134 E-Mail: ankitasinha051@gmail.com

ABSTRACT
The emergence of Deep Learning (DL) technologies, coupled with advancements in computing power, has significantly benefited the field of medical image diagnosis. DL technologies have helped in many fields of medical imaging like segmentation, feature extraction, and classification. These DL technologies have been published widely in various medical imaging analyzing methods. In this paper, we review deep learning based Medical image harmonization techniques.

Keywords:
Deep Learning (DL), Medical Imaging, Harmonization.

1. INTRODUCTION
Present days are very much an era of Artificial intelligence. With the increase and availability of big data its subareas machine learning and DL has affected everyone's life. DL has wide applications in many scientific fields such as computer vision, natural language processing, audio & speech recognition and bioinformatics(Deng L, 2013). DL has its root in the evolution of theory of Perceptron, first generation multi-layer perceptron way back in 1950s and 1960s. However, its widespread adoption truly gained momentum following the 2012 large-scale image categorization challenge, which conclusively demonstrated the superiority of Convolutional Neural Networks (CNNs) on the ImageNet dataset (Krizhevsky et al., 2017). In the last decade, with the advancement of computing power and availability of large scale datasets, DL has penetrated its application in many real world applications such as advertisements and helped in solving real life problems. One of them is in the field of medical diagnosis. With the growth in digitization, a huge amount of medical data is generated. Medical applications use imaging modalities like X-ray, ultrasound, computed tomography(CT) scan, magnetic resonance imaging (MRI) and few others. DL is primarily based on multiple layers of neural networks resembling human brain. Medical imaging has several traits that correlate with the nature of DL solutions (S. K. Zhou et al., 2021). DL helps in many important medical imaging techniques like classification, segmentation, image reconstruction, image enhancement and dataset augmentation.
Medical image data is often collected and obtained in non-standardized settings. The absence of standardized acquisition protocols leads to variations in equipment and scanning parameters, contributing to a phenomenon known as "distribution drift" (S. K. Zhou et al., 2021). This phenomenon introduces confounding effects caused by non-biological sources of variability in the data, primarily arising from discrepancies in image acquisition hardware and protocols. To facilitate improved analysis of these data, harmonization is essential.
Multicenter data harmonization refers to the application of mathematical and statistical methodologies aimed at mitigating undesired variability across multiple research sites called scanner effect or batch effect, all the while preserving the essential biological information inherent within the data (Marzi et al., 2024).

This article provides a survey on different DL harmonization techniques. The article is structured as follows: Section 2 provides an overview of deep learning methods. Section 3 provides an overview of different harmonization techniques based on DL solutions for MRI images. Section 4 provides evaluation of reviewed harmonization techniques. Section 5 provides conclusion and future work.

2. DEEP LEANING

Deep Learning encompasses a collection of computationally intensive models, which are fully connected acyclic multi-layer neural networks. Deep learning methodologies applied within the domain of medical imaging exhibit versatile capabilities, encompassing functions ranging from the acquisition of medical images to the detection and characterization of pathological conditions depicted within these images. More precisely, these methodologies are utilized not solely to augment the fidelity of images acquired across diverse modalities but also to facilitate the discernment of pathological indicators within such images in a manner that is both effective and resource-efficient. For example, convolutional neural networks (CNNs) find utility in the reconstruction of images derived from MRI scanners, thereby augmenting image resolution and enabling improved visualization of potential pathological features (Zhang & Qie, 2023). CNNs can also do image classification, wherein discerning the presence or absence of particular anatomical or pathological attributes. CNNS can also help in image segmentation (Shelhamer et al., 2017), wherein it identifies and outlines specific structures or regions within the image. Deep learning methodologies also contribute significantly to image enhancement and dataset augmentation (Chen et al., 2022) within the field by generating synthetic images. It is achieved with the help of Generative Adversarial Networks (GANs)(Goodfellow et al., 2020). The quantity of training samples utilized in medical image-based diagnostic and therapeutic models is expanding alongside the advancements in deep learning. Notably, Generative Adversarial Networks (GANs), renowned for their remarkable image generation capabilities and widespread application in data augmentation, have garnered attention within the realm of medical image processing. Following subsections briefly describes different deep learning technologies.

2.1 Feed Forward Neural Network.

It is a multi-layer perceptron. Deep learning approximates a function with the help of some hidden layers. Gradient descent is a well-known optimisation method used to train the feed forward neural network. For approximating the function, noisy estimates are given in the form of training datasets. Neural networks are usually trained by minimising the loss function.

2.2 Convolution Neural Network

A general model of Convolutional Neural Networks (CNNs) comprises four fundamental processes: convolutional layers, pooling layers, fully connected layers, and activation functions.

Convolutional Layer: Convolutional layer's primary feature is feature extraction. Each neuron within the subsequent layer establishes connections with specific neurons in the preceding layer, defining a local correlation known as the receptive field (Neubauer, 1998). Through this mechanism, local features inherent in the input image are extracted.

Pooling Layer: The pooling layer functions to merge semantically akin features into one while simultaneously diminishing the dimensionality of the representation (Lecun et al., 2015). The primary benefit of employing pooling techniques lies in their notable reduction of trainable parameters while also imparting translation invariance (Y. Zhou et al., 2016). Among various pooling methods such as average pooling and max-pooling, the latter is widely favored for its substantial reduction in map size (Lee et al., 2017).

Fully Connected Layer: Fully connected layers enable the integration of local features across spatial dimensions, thereby facilitating the emergence of global patterns essential for image classification within neural networks. These layers are instrumental in executing classification tasks by leveraging the features extracted from preceding layers.

Activation Function: The activation function typically serves as the final layer in a Convolutional Neural Network (CNN) and is primarily employed for classification purposes.

2.3 Recurrent Neural Networks:

Recurrent Neural Networks (RNNs) represent a class of artificial neural networks adept at sequential data processing within the domain of deep learning. In contrast to traditional feed forward neural networks, Recurrent Neural Networks (RNNs) possess a notable capacity to maintain internal states, thereby preserving contextual information across sequential input processing epochs. This inherent memory feature facilitates the retention of learned patterns across successive time steps, empowering RNNs to effectively capture temporal dependencies within sequential data streams (Cho et al., 2014).

2.4Auto-encoder

Auto-encoders were initially conceptualized in (David E. Rumelhart, James L. McClelland, 1986) as neural network architectures designed to learn representations of input data through the process of reconstruction. Functionally, an auto-encoder endeavors to generate output that closely resembles its input, thereby serving as an algorithmic tool for image synthesis. Conceptually, auto-encoders extend the principles of Principal Component Analysis (PCA) (Plaut, 2018), facilitating the discovery of latent representations underlying input data while minimizing information loss during reconstruction.

2.5 Generative Adversarial Network (GAN):

With the recent research in Generative Adversarial Networks (GANs), realistic-looking natural image generations have become a possibility(Goodfellow et al., 2020). Generative Adversarial Networks (GANs) find widespread utility in medical imaging, particularly for tasks such as data augmentation, image reconstruction, and image-to-image translation. The training paradigm of GANs is structured as a two-agent adversarial game, featuring a generator (G) and a discriminator (D). Within this framework, the generator seeks to produce images possessing high fidelity to real counterparts, whereas the discriminator endeavors to distinguish genuine images from those synthesized by the generator. This adversarial interplay between the generator and discriminator drives the iterative refinement of generated images towards enhanced realism, facilitating the attainment of desired performance in medical imaging applications.

3. HARMONIZATION METHODS

In recent years, a diverse array of harmonization methodologies has emerged as potent and adaptable mechanisms for mitigating confounding effects attributable to site, scanner, or protocol disparities, thereby maintaining the inherent biological information embedded within images. Traditional post-processing techniques, including global scaling, and functional normalization (J. Fortin et al., 2014), have demonstrated efficacy in diminishing the impact of site or scanner-related biases. Nevertheless, they have proven inadequate in adequately addressing the spatial heterogeneity inherent in site effects (J. P. Fortin et al., 2017).

Harmonization methods are broadly classified into statistical and deep learning based methods. In Statistical techniques, Intensity normalization is used in White Stripe (Shinohara et al., 2014), which systematically standardizes the normal appearing white matter (NAWM) intensity distributions. Another work which uses Intensity normalization is used in **M**ultisite **i**mage harmonization by **c**umulative distribution function **a**lignment (MICA) (Wrobel et al., 2020).This technique entails aligning voxel intensity cumulative distribution functions (CDFs) to achieve harmonization. This approach entails the estimation of nonlinear, monotonically increasing transformations applied to voxel intensity values within a single scan. These transformations are designed to align the cumulative distribution function (CDF) of intensity values precisely with that of a specified reference scan, denoted as the "target" scan. In a different statistical methodology for addressing batch effects, Removal of Artificial Voxel Effect by Linear Regression (RAVEL) (J. P. Fortin et al., 2016) is

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 2, April : 2024

introduced as a proposed technique. It estimates the latent factors of unwanted variations common to all voxels. In other batch effect adjustment method, Combat (J. P. Fortin et al., 2017) is proposed. It identifies batch specific transformation to express all data in common representation.

3.1 Deep Learning based harmonization methods

In recent years, a plethora of deep learning methodologies have emerged as potent and versatile tools for the harmonization of MRI images. These deep learning methodologies are mainly base on different version of GAN based architectures, auto-encoders, U-net architectures. A review of few deep learning based harmonization methodologies is done in this paper.

3.1.1 Surface-to-Surface GAN

The Surface-to-Surface GAN (S2SGAN) (Zhao et al., 2019) is a variant of CycleGAN (Zhu et al., 2017) that employs spherical U-Net layers instead of conventional convolution layers. This adaptation is designed specifically for performing harmonization tasks on subject-wise cortical thicknesses projected onto a spherical surface. The Surface-to-Surface GAN (S2SGAN) model operates on the task of harmonization from one surface, denoted as X, to another surface, denoted as Y, thereby constituting a surface-to-surface translation endeavor. In this context, the primary objective of harmonization is to acquire a mapping, $G_X{:}X\rightarrow Y$, such that the distribution of $G_X(X)$ closely resembles that of Y. Given the under-constrained nature of this mapping, an additional goal of harmonization is to preserve biological variance. To achieve this, the model employs the inverse mapping $G_Y{:}Y\rightarrow X$ alongside the cycle consistency loss, enforcing $G_Y(G_X(X))\approx X$ and vice versa.

Moreover, to ensure structural consistency between the original surface thickness maps and the generated maps, the model incorporates a correlation coefficient loss (Hu, Chen, et al., 2023).

3.1.2 DeepHarmony

Deep-Harmony (Dewey et al., 2019) presents a contrast harmonization method utilizing a convolutional U-Net architecture aimed at achieving consistent contrast in medical imaging. The methodology is designed to harmonize images directly, leveraging multiple contrasts (e.g., T1-weighted, FLAIR, T2-weighted/proton density) obtained from each subject under differing acquisition protocols. Through a "many-to-many" reconstruction strategy, complementary information across contrasts is integrated to reconstruct corresponding contrasts under different protocols.

Key modifications are made to the architecture to optimize harmonization performance compared to standard U-Net implementations. Notably, a final concatenation operation is introduced between input contrasts and the concluding feature map, guiding the network to enhance input contrasts rather than fully reconstructing target contrasts. Strided convolution and deconvolution are also implemented for downsampling and upsampling feature maps, respectively, enhancing the network's ability to harmonize images across multiple contrasts.

By prioritizing the transformation of input data to accurately reconstruct desired output, the network improves efficiency and effectiveness in harmonization, minimizing the need for complete reconstruction of reference contrasts.

3.1.3 CALAMITI

The Contrast Anatomy Learning and Analysis for MR Intensity Translation and Integration (CALAMITI) (Zuo et al., 2021) method represents an unsupervised harmonization approach that amalgamates the advantages of both image-to-image translation and unsupervised domain adaptation techniques. It is based on a conditional variational auto-encoder model.

CALAMITI represents an advanced, theoretically grounded, unsupervised harmonization approach rooted in information bottleneck (IB) principles. It endeavors to learn a global, disentangled latent space encompassing anatomical and contrast information, facilitating seamless adaptation to new testing sites solely based on incoming data. Notably, this work addresses four key challenges in unsupervised harmonization.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 2, April : 2024

Firstly, by leveraging intra-site paired data, CALAMITI tackles unsupervised image-to-image translation (IIT) tasks through a supervised approach, obviating the need for additional constraints such as cycle-consistency and achieving superior pixel-to-pixel regularization.
Secondly, it adopts a unified architecture for multi-site harmonization, ensuring that the model size remains independent of the number of sites involved.
Thirdly, CALAMITI fosters a consistent anatomical description across all training data, thereby establishing a global latent space that encapsulates anatomical variations.
Finally, the method exhibits adaptability to new sites without necessitating retraining on the original dataset. Additionally, theoretical underpinnings are provided, elucidating the disentangled nature of the latent space through IB theory (Zuo et al., 2021).

3.1.4 Imunity

ImUnity (Cackowski et al., 2023), a deep learning-based methodology, extends previous techniques to offer a swift and flexible harmonization solution. It employs a self-supervised Variational AutoEncoder (VAE-GAN) architecture to generate "corrected" magnetic resonance (MR) images applicable to varied population imaging studies. To mitigate the requirement for mobile subjects or diverse MR sequences in the database, ImUnity utilizes multiple slices from the same individual during training, coupled with randomized image contrast transformations.
Furthermore, it incorporates a mechanism to mitigate center bias through a confusion module linked to its bottleneck. Optionally, a biological module is employed to ensure the preservation of clinical features within the latent space. Once trained, this architecture enables the harmonization of data from new sites or scanners without necessitating fine-tuning. Moreover, it supports estimates for multiple target sites, enabling users to select MR image reconstructions tailored to the desired target domain, be it a site or a scanner.
The generator network receives two 2-D structural images in identical orientations as input. The initial image (S1) is processed by the first convolutional neural network (CNN), exclusively utilizing convolutional filters, to encode the 'anatomical' information. This design choice aims to maintain spatial information integrity throughout the encoding process. The second image (S_2), distinct from the first image (S_1), is randomly sampled from a different brain region, providing an initial source of 'contrast' information. Consequently, S_1 and S_2 exhibit divergent anatomical characteristics due to their disparate locations within the brain, while sharing similar contrast attributes stemming from identical scanning protocols. The contrast of the second image, S_2, is adjusted using a gamma function (or exponential correction). The gamma parameter is randomly sampled from a uniform distribution ranging between 0.5 and 1.5 for each new input 2-D slice. The altered S^{γ}_2 slice serves as the input to a secondary convolutional neural network (CNN) tasked with encoding the 'contrast information'. Subsequently, a dense layer is employed to diminish spatial information within the encoded representation.
Following encoding, the distinct representations of S_1 and S^{γ}_2 are concatenated to form a latent space representation. This latent space representation is then decoded using transposed convolutional filters to generate the output S^{γ}_1.
In this model, the separation of content information from style is accomplished through a self-supervised process, eliminating the need for additional imaging contrasts. Image harmonization is accomplished by feeding reference batch slices to the style encoder and source batch slices to the content encoder. If unseen batches exhibit sufficient similarity to the training batches, allowing the content encoder to effectively embed slices from unseen batches, the model can readily adapt to these settings (Hu, Chen, et al., 2023).

3.1.5 MISPEL

Multi-scanner Image Harmonization via Structure Preserving Embedding Learning (MISPEL) (Torbati et al., 2023) aims to achieve harmonization across multiple batches, where the number of batches, denoted as *m*, can exceed two. This is accomplished by employing a collection of m batch-specific convolutional autoencoders, trained through a two-step algorithm. Importantly, the encoders

are constructed as deep neural networks, while the decoders execute a linear combination of the latent-space representations.
In the initial step, MISPEL facilitates the training of each batch-specific encoder to encode slices from its respective batch into a unified latent space. Following this, the associated decoder is trained to reconstruct slices in the stylistic manner of its batch utilizing these latent-space representations. This procedure involves individually training each batch-specific autoencoder in a self-supervised fashion, employing a reconstruction loss. Furthermore, MISPEL ensures a shared latent space across all autoencoders by integrating a representation similarity loss, which penalizes excessive divergence among latent-space representations.
In the second step, the encoders are held constant, and only the decoders are updated. This ensures that all decoders produce harmonized output slices with similar characteristics, while also maintaining similarity between the outputs and the input slices. (Hu, Chen, et al., 2023)
3.1.6 MURD
The Multi-site Unsupervised Representation Disentangler (MURD) (Liu & Yap, 2024)is a harmonization method grounded in deep neural networks, eliminating the dependency on traveling human phantom data. This technique operates by disentangling site-specific appearance characteristics and site-invariant anatomical features from images acquired across multiple sites. Subsequently, the disentangled information is utilized to generate images of each subject for any target site.
MURD decomposes images into two distinct components: anatomical content, which remains consistent across sites, and appearance style, encompassing attributes such as intensity and contrast, which vary between sites. Harmonized images are synthesized by integrating the content of an image with styles specific to the respective sites.
This encoding process entails employing two encoders: a content encoder capturing anatomical structures shared across sites, and a style encoder capturing site-specific style information. An image harmonized for a specific site is produced by a generator that merges the extracted content with the style characteristic of the target site. The target style can be specified either through a reference image from the target site or by a randomly generated style code from a site-specific style generator. The latter method allows for the creation of multiple visual appearances that capture natural variations in style across each site.
MURD is trained with tailored loss functions aimed at promoting comprehensive representation disentanglement while maximizing the preservation of structural details within the harmonized images.
3.1.7 DeepCombat
DeepCombat (Hu, Lucas, et al., 2023) is a deep learning harmonization approach based on a conditional variational autoencoder (CVAE) architecture, integrated with the ComBat harmonization model. It aims to learn and eliminate subject-level batch effects by considering the multivariate relationships among features, presenting a novel perspective within statistically-motivated deep learning harmonization methodologies. DeepCombat offers the capability to address complex, non-linear, and multivariate batch effects within raw data, minimizing the detectability of such effects through highly multivariate techniques.
The DeepCombat method involves three primary steps: normalization, CVAE training, and harmonization. Initially, the normalization step transforms the raw data to facilitate quicker convergence during CVAE training. Subsequently, the CVAE endeavors to learn a latent space representation of the input data, enriched with subject-specific information while minimizing batch effects. In this phase, the CVAE learns to utilize this latent space representation, along with explicit batch and biological information, for data reconstruction.
Following reconstruction, imperfections and residual batch effects are addressed through harmonization using ComBat. The CVAE decoder then employs this harmonized latent space, coupled with reference batch covariates, to generate harmonized subject-specific means. Finally, harmonized residuals are combined with these means to obtain the ultimate harmonized data.

In summary, DeepCombat segregates batch effects into three distinct components—the latent space, the CVAE decoder, and the reconstruction residuals. Each component undergoes individual harmonization before amalgamation to produce the final DeepCombat-harmonized data.

4. EVALUATION

Successful harmonization methods in medical imaging must fulfill two primary objectives: (1) effectively mitigating undesired variations arising from site-specific factors and disparities in imaging protocols, and (2) preserving inherent biological variability among subjects. S2SGAN utilizes mean absolute error (MAE) and peak signal-to-noise ratio (PSNR) metrics for performance evaluation, demonstrating superior results over Combat in both MAE and PSNR for cortical thickness. DeepHarmony employs Mean Structured Similarity Index (SSIM) and MAE, averaged across all four contrasts, as validation metrics, showcasing significant enhancement in harmonization across all measurements. CALAMITI achieves state-of-the-art harmonization performance visually and in terms of SSIM and PSNR. However, limitations exist, such as the necessity of intra-site paired images during training, potentially restricting applications like pediatric data acquisition where acquiring multi-contrast images is impractical. Moreover, experiments solely employ paired T1-w and T2-w images, although extending the method to include additional contrast MR images like fluid-attenuated inversion recovery (FLAIR) could enhance disentanglement. Thirdly, experiments on multiple sclerosis (MS) patients reveal unsatisfactory harmonization results in white matter (WM) lesion areas. ImUnity employs SSIM metrics and Support Vector Machine (SVM) classifiers for harmonization evaluation, albeit its suitability for small sample size scenarios may be limited.

MISPEL adopts three evaluation criteria: visual quality, image similarity, and volumetric similarity. Image similarity is assessed using the mean structural similarity index measure (SSIM), indicating improved similarity across harmonized scans. Volumetric similarity is evaluated through the Dice similarity coefficient (DSC), indicating enhanced tissue segmentation similarity.

MURD assesses harmonization efficacy using mean absolute error (MAE), multi-scale structural similarity (MS-SSIM), and peak signal-to-noise ratio (PSNR) on T1-weighted and T2-weighted images from the traveling human phantom dataset. MURD effectively harmonizes MR images by mitigating non-biological site differences while preserving anatomical details.

Evaluation of DeepCombat involves qualitative visualization, statistical testing, and machine learning experiments.

5. CONCLUSION AND FUTURE WORK

Recent advancements in deep learning architectures have demonstrated their potential to enhance diagnostic precision in medical imaging across various domains including pathology and neuroimaging. However, the effective deployment of deep learning models necessitates access to substantial volumes of data. To facilitate more robust data analysis, it is imperative to address batch or scanner effects through harmonization techniques.

This review paper systematically examines multiple deep learning-based harmonization methodologies within the realm of medical imaging. It elucidates their fundamental architectural principles and operational methodologies, while also evaluating their performance metrics and identifying inherent limitations. Each method endeavours to mitigate batch or scanner effects through distinct approaches. Furthermore, the review discusses a hybrid method that integrates statistical and deep learning techniques.

It is essential for methodologists to strive towards establishing a standardized harmonization approach that comprehensively addresses all relevant aspects. Such an endeavour would furnish end-users with a reliable tool for data analysis and prediction in medical imaging applications.

References

1. Cackowski, S., Barbier, E. L., Dojat, M., & Christen, T. (2023). ImUnity: A generalizable VAE-GAN solution for multicenter MR image harmonization. *Medical Image Analysis*, *88*.

https://doi.org/10.1016/j.media.2023.102799
2. Chen, Y., Yang, X. H., Wei, Z., Heidari, A. A., Zheng, N., Li, Z., Chen, H., Hu, H., Zhou, Q., & Guan, Q. (2022). Generative Adversarial Networks in Medical Image augmentation: A review. *Computers in Biology and Medicine*, *144*(March), 105382. https://doi.org/10.1016/j.compbiomed.2022.105382
3. Cho, K., van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder–decoder approaches. *Proceedings of SSST 2014 - 8th Workshop on Syntax, Semantics and Structure in Statistical Translation*, 103–111. https://doi.org/10.3115/v1/w14-4012
4. David E. Rumelhart, James L. McClelland, P. R. G. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations* (Vol. 1, Issue January 1986).
5. Deng L, Y. D. (2013). DL: methods and applications. Foundations and Trends® in Signal Processing. In *Revista Europea de Estudios Latinoamericanos y del Caribe* (Vol. 95, pp. 71–95).
6. Dewey, B. E., Zhao, C., Reinhold, J. C., Carass, A., Fitzgerald, K. C., Sotirchos, E. S., Saidha, S., Oh, J., Pham, D. L., Calabresi, P. A., van Zijl, P. C. M., & Prince, J. L. (2019). DeepHarmony: A deep learning approach to contrast harmonization across scanner changes. *Magnetic Resonance Imaging*, *64*, 160–170. https://doi.org/10.1016/j.mri.2019.05.041
7. Fortin, J., Lemire, M., Brent, W., Hudson, T. J., Fertig, E. J., & Greenwood, C. M. T. (2014). *Functional normalization of 450k methylation array data improves replication in large cancer studies*. 1–42.
8. Fortin, J. P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., Schultz, R. T., Verma, R., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, *161*, 149–170. https://doi.org/10.1016/j.neuroimage.2017.08.047
9. Fortin, J. P., Sweeney, E. M., Muschelli, J., Crainiceanu, C. M., & Shinohara, R. T. (2016). Removing inter-subject technical variability in magnetic resonance imaging studies. *NeuroImage*, *132*(May 2016), 198–212. https://doi.org/10.1016/j.neuroimage.2016.02.036
10. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139–144. https://doi.org/10.1145/3422622
11. Hu, F., Chen, A. A., Horng, H., Bashyam, V., Davatzikos, C., Alexander-Bloch, A., Li, M., Shou, H., Satterthwaite, T. D., Yu, M., & Shinohara, R. T. (2023). Image harmonization: A review of statistical and deep learning methods for removing batch effects and evaluation metrics for effective harmonization. *NeuroImage*, *274*(February), 120125. https://doi.org/10.1016/j.neuroimage.2023.120125
12. Hu, F., Lucas, A., Chen, A. A., Coleman, K., Horng, H., Ng, R. W. S., Tustison, N. J., Davis, K. A., Shou, H., Li, M., Shinohara, R. T., & Alzheimer's Disease Neuroimaging Initiative. (2023). DeepComBat: A Statistically Motivated, Hyperparameter-Robust, Deep Learning Approach to Harmonization of Neuroimaging Data. *BioRxiv : The Preprint Server for Biology*. https://doi.org/10.1101/2023.04.24.537396
13. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386
14. Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539
15. Lee, K. B., Cheon, S., & Kim, C. O. (2017). A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing*, *30*(2), 135–142. https://doi.org/10.1109/TSM.2017.2676245
16. Liu, S., & Yap, P.-T. (2024). Learning multi-site harmonization of magnetic resonance images without traveling human phantoms. *Communications Engineering*, *3*(1), 1–10.

https://doi.org/10.1038/s44172-023-00140-w
17. Marzi, C., Giannelli, M., Barucci, A., Tessa, C., Mascalchi, M., & Diciotti, S. (2024). Efficacy of MRI data harmonization in the age of machine learning: a multicenter study across 36 datasets. *Scientific Data*, *11*(1), 1–27. https://doi.org/10.1038/s41597-023-02421-7
18. Neubauer, C. (1998). Evaluation of convolutional neural networks for visual recognition. *IEEE Transactions on Neural Networks*, *9*(4), 685–696. https://doi.org/10.1109/72.701181
19. Plaut, E. (2018). *From Principal Subspaces to Principal Components with Linear Autoencoders*. 1–6. http://arxiv.org/abs/1804.10253
20. Shelhamer, E., Long, J., & Darrell, T. (2017). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(4), 640–651. https://doi.org/10.1109/TPAMI.2016.2572683
21. Shinohara, R. T., Sweeney, E. M., Goldsmith, J., Shiee, N., Mateen, F. J., Calabresi, P. A., Jarso, S., Pham, D. L., Reich, D. S., & Crainiceanu, C. M. (2014). Statistical normalization techniques for magnetic resonance imaging. *NeuroImage: Clinical*, *6*, 9–19. https://doi.org/10.1016/j.nicl.2014.08.008
22. Torbati, M. E., Minhas, D. S., Laymon, C. M., Maillard, P., Wilson, J. D., Chen, C. Le, Crainiceanu, C. M., DeCarli, C. S., Hwang, S. J., & Tudorascu, D. L. (2023). MISPEL: A supervised deep learning harmonization method for multi-scanner neuroimaging data. *Medical Image Analysis*, *89*(June), 102926. https://doi.org/10.1016/j.media.2023.102926
23. Wrobel, J., Martin, M. L., Bakshi, R., Calabresi, P. A., Elliot, M., Roalf, D., Gur, R. C., Gur, R. E., Henry, R. G., Nair, G., Oh, J., Papinutto, N., Pelletier, D., Reich, D. S., Rooney, W. D., Satterthwaite, T. D., Stern, W., Prabhakaran, K., Sicotte, N. L., … Goldsmith, J. (2020). Intensity warping for multisite MRI harmonization. *NeuroImage*, *223*(December 2020), 1–18. https://doi.org/10.1016/j.neuroimage.2020.117242
24. Zhang, H., & Qie, Y. (2023). Applying Deep Learning to Medical Imaging: A Review. *Applied Sciences (Switzerland)*, *13*(18). https://doi.org/10.3390/app131810521
25. Zhao, F., Wu, Z., Wang, L., Lin, W., Xia, S., Shen, D., & Li, G. (2019). Harmonization of infant cortical thickness using surface-to-surface cycle-consistent adversarial networks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11767 LNCS* (pp. 475–483). https://doi.org/10.1007/978-3-030-32251-9_52
26. Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., Van Ginneken, B., Madabhushi, A., Prince, J. L., Rueckert, D., & Summers, R. M. (2021). A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies with Progress Highlights, and Future Promises. *Proceedings of the IEEE*, *109*(5), 820–838. https://doi.org/10.1109/JPROC.2021.3054390
27. Zhou, Y., Wang, H., Xu, F., & Jin, Y. Q. (2016). Polarimetric SAR Image Classification Using Deep Convolutional Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, *13*(12), 1935–1939. https://doi.org/10.1109/LGRS.2016.2618840
28. Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. *Proceedings of the IEEE International Conference on Computer Vision*, *2017-Octob*, 2242–2251. https://doi.org/10.1109/ICCV.2017.244
29. Zuo, L., Dewey, B. E., Carass, A., Liu, Y., He, Y., Calabresi, P. A., & Prince, J. L. (2021). Information-Based Disentangled Representation Learning for Unsupervised MR Harmonization. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12729 LNCS*, 346–359. https://doi.org/10.1007/978-3-030-78191-0_27