



**HAL**  
open science

# MixUp brain-cortical augmentations in self-supervised learning

Corentin Ambroise, Vincent Frouin, Benoit Dufumier, Edouard Duchesnay,  
Antoine Grigis

► **To cite this version:**

Corentin Ambroise, Vincent Frouin, Benoit Dufumier, Edouard Duchesnay, Antoine Grigis. MixUp brain-cortical augmentations in self-supervised learning. *Medical Image Computing and Computer Assisted Intervention*, Oct 2023, Vancouver, Canada. pp.102-111, 10.1007/978-3-031-44858-4\_10. hal-04471607

**HAL Id: hal-04471607**

**<https://hal.science/hal-04471607v1>**

Submitted on 22 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# MixUp Brain-Cortical Augmentations in Self-supervised Learning

Corentin Ambroise<sup>(✉)</sup>, Vincent Frouin, Benoit Dufumier, Edouard Duchesnay,  
and Antoine Grigis

Université Paris-Saclay, CEA, NeuroSpin, 91191 Gif-sur-Yvette, France  
[corentin.ambroise9132@gmail.com](mailto:corentin.ambroise9132@gmail.com)

**Abstract.** Learning biological markers for a specific brain pathology is often impaired by the size of the dataset. With the advent of large open datasets in the general population, new learning strategies have emerged. In particular, deep representation learning consists of training a model via pretext tasks that can be used to solve downstream clinical problems of interest. More recently, self-supervised learning provides a rich framework for learning representations by contrasting transformed samples. These methods rely on carefully designed data manipulation to create semantically similar but syntactically different samples. In parallel, domain-specific architectures such as spherical convolutional neural networks can learn from cortical brain measures in order to reveal original biomarkers. Unfortunately, only a few surface-based augmentations exist, and none of them have been applied in a self-supervised learning setting. We perform experiments on two open source datasets: Big Healthy Brain and Healthy Brain Network. We propose new augmentations for the cortical brain: baseline augmentations adapted from classical ones for training convolutional neural networks, typically on natural images, and new augmentations called MixUp. The results suggest that surface-based self-supervised learning performs comparably to supervised baselines, but generalizes better to different tasks and datasets. In addition, the learned representations are improved by the proposed MixUp augmentations. The code is available on GitHub ([https://github.com/neurospin-projects/2022\\_cambroise\\_surfaugment](https://github.com/neurospin-projects/2022_cambroise_surfaugment)).

**Keywords:** Data augmentation · Spherical convolutional neural networks · Self-supervised learning · Brain structural MRI

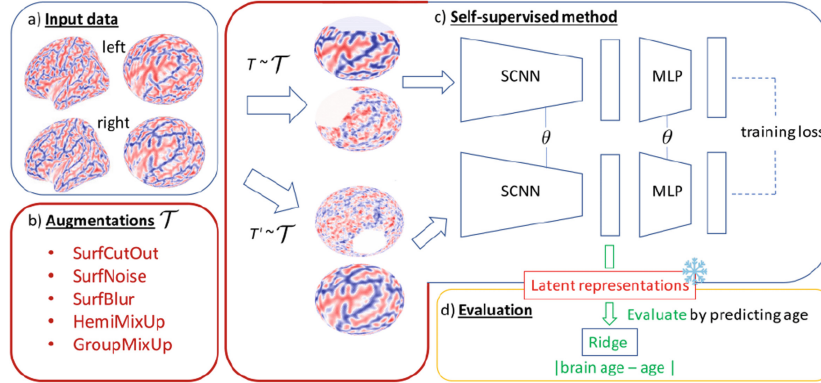
## 1 Introduction

Data-driven studies of brain pathologies are often hampered by the scarcity of available data, leading to potential failures in the discovery of statistically significant biomarkers. Key factors include recruitment of rare disease patients and

---

**Supplementary Information** The online version contains supplementary material available at [https://doi.org/10.1007/978-3-031-44858-4\\_10](https://doi.org/10.1007/978-3-031-44858-4_10).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023  
A. Abdulkadir et al. (Eds.): MLCN 2023, LNCS 14312, pp. 102–111, 2023.  
[https://doi.org/10.1007/978-3-031-44858-4\\_10](https://doi.org/10.1007/978-3-031-44858-4_10)



**Fig. 1.** Overview of the proposed evaluation framework for spherical augmentations: a) input cortical measures (here curvature) inflated to a sphere, b) a set of adapted and domain-specific augmentations  $\mathcal{T}$  that allow the generation of augmented cortical measures, c) a self-supervised method with parameters  $\theta$  consisting of a spherical Convolutional Neural Network (CNN) and a Multi-Layer Perceptron (MLP) projector, and d) evaluation of the model’s frozen representations using a linear predictor (here the mean absolute deviation between linearly predicted brain age and true age).

acquisition costs. Many research efforts have attempted to address this challenge [12]. Transfer learning has become a promising solution with the advent of large cohorts such as Big Healthy Brain (BHB) [11]. Transfer learning consists of training a Neural Network (NN) with pretext tasks on a large dataset. The trained NN is then fine-tuned on a smaller, application-specific dataset. However, transfer learning for medical imaging is still in its early stages. Interestingly, there is a consensus that using a natural image dataset may not lead to the best transfer strategy [2, 21]. In recent years, several training schemes have been proposed for learning “universal” data representations [3]. The goal is to summarize as much of the semantic information as possible. Among the most promising approaches are self-supervised schemes that can provide good NN initialization for transfer learning [4, 16, 29]. Specifically, contrastive learning uses data augmentation to structure the learned latent space [5, 17, 19]. Therefore, such techniques rely heavily on the data augmentation [4, 5, 16, 28]. Currently, data augmentation for medical imaging is only available for image data defined on a regular rectangular grid.

In this work, we focus on domain-specific architectures called Spherical Convolutional Neural Networks (SCNNs). SCNNs have the potential to discover novel biomarkers from cortical brain measures. The bottleneck is the definition of convolution strategies adapted to graph or spherical topologies. Several strategies have been proposed in the literature [6, 8, 9, 18, 20, 23, 25, 32]. In neuroimaging, cortical measures are usually available on the left and right brain hemisphere surfaces of each individual. These surfaces can be inflated to icosahedral spheres [13]. Leveraging the regular and consistent geometric structure of this spherical mesh on which the cortical surface is mapped, Zhao *et al.* defined the Direct Neighbor (DiNe) con-

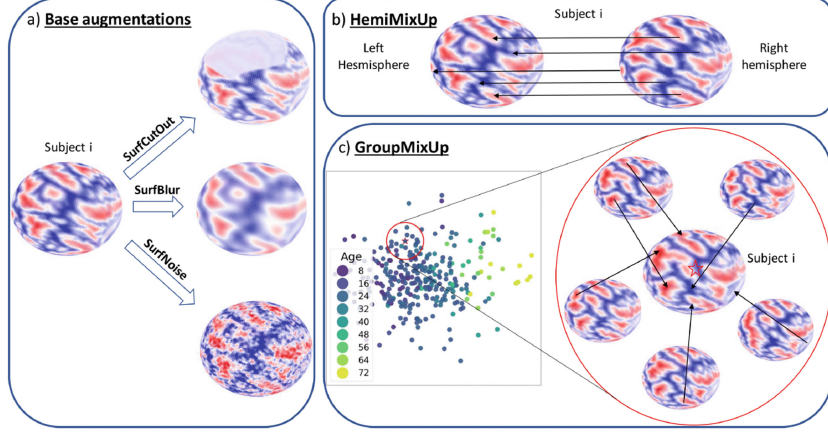
volution filter on the sphere [32]. The DiNe convolution filter is analogous to the standard convolution on the image grid, but defines its neighbors as concentric rings of oriented vertices (see Supplemental S1).

To train SCNNs using contrastive learning, we introduce three baseline and two original augmentations specifically designed for the brain cortical surfaces (Fig. 1). The three baseline augmentations are directly inspired by natural image transformations: the SurfCutOut involves cutting out surface patches, the SurfNoise adds Gaussian noise at each vertex, and the SurfBlur applies Gaussian blur. The proposed MixUp augmentations build on the original idea of randomly selecting some cortical measures and replacing them with realistic corrupted samples: the HemiMixUp exploits the symmetry of the brain and permutes measures between hemispheres of the same individual, and the GroupMixUp bootstraps vertex-based measures from a group of similar individuals. In this work, we illustrate how these augmentations can fit into the well-known SimCLR self-supervised scheme [5]. We provide a comprehensive analysis showing that self-supervised learning with SCNNs and the proposed augmentations shows similar predictive performance as supervised SCNNs for predicting age, sex and a cognitive phenotype from structural brain MRI. Furthermore, we show that the MixUp augmentations improve the learned representations and the generalization performance of the self-supervised model.

## 2 Methods

### 2.1 Baseline Brain Cortical Augmentations

All augmentations defined for natural and medical images are not directly applicable to the cortical surface. In a self-supervised scheme, an effective augmentation must reflect invariances that we want to enforce in our representation. It is not a requirement that these augmentations produce realistic samples. Their goal is to provide synthetic contrastive hard prediction tasks [5, 24]. They must also be computationally efficient. For example, geometric transformations such as cropping, flipping, or jittering cannot be applied to cortical measures. We could use small rotations as proposed in [31], but such an augmentation is not computationally efficient due to multiple interpolations and less effective in nonlinear registration cases. We adapt three baseline domain-specific transformations consisting of cutting out surface patches (SurfCutOut), adding Gaussian noise at each vertex (SurfNoise), and Gaussian blurring (SurfBlur). Specifically, the SurfCutOut sets an adaptive neighborhood around a random vertex to zero. The neighborhood is defined by  $R$  concentric rings of vertices (for a definition of a ring, see the previously introduced DiNe operator). On structural MRI images, a cutout strategy has proven its efficiency in a similar contrastive learning setting [10]. Then, the SurfNoise adds a Gaussian white noise with standard deviation  $\sigma_1$  (to weight the signal-to-noise ratio), and the SurfBlur smooths the data by applying a Gaussian kernel with standard deviation  $\sigma_2$  (which controls the spatial extent expressed in rings).



**Fig. 2.** Illustration of the considered cortical augmentations: a) the three baselines, b) the proposed HemiMixUp, and c) the proposed GroupMixUp brain cortical augmentations. In the GroupMixUp, groups are defined from the reduced input data using PCA embeddings colored by age. The arrows in b) and c) represent how the cortical measures are modified and are explained in more detail in Supplemental S2.

## 2.2 Proposed MixUp Brain Cortical Augmentations

Here, we assume that the structural measures across the cortical surface have a vertex-to-vertex correspondence for both hemispheres. We propose to randomly select vertices and their associated cortical measures and to replace them with noisy realistic ones. A similar approach has been proposed by Yoon et al. [27] for tabular data, and a comparable augmentation has been used for natural images in supervised contexts by mixing up labels [30]. All augmentations are applied on a cortical measure and hemisphere basis. A corrupted version  $\tilde{\mathbf{x}}$  of a cortical measure  $\mathbf{x} \in \mathbb{R}^P$ , where  $P$  is the number of vertices, is generated as follows:

$$\tilde{\mathbf{x}} = \mathbf{m} \odot \bar{\mathbf{x}} + (1 - \mathbf{m}) \odot \mathbf{x} \quad (1)$$

where  $\odot$  is the point-wise multiplication operator,  $\bar{\mathbf{x}}$  is a noisy sample to be defined, and  $\mathbf{m} \in \{0, 1\}^P$  is a binary random mask. In our case  $\mathbf{x} \in \{x_1, \dots, x_N\}$ , where  $N$  is the number of subjects. The proposed HemiMixUp and GroupMixUp augmentations offer different ways to construct  $\bar{\mathbf{x}}$  (Fig. 2). The mask  $\mathbf{m}$  is generated by drawing binary random numbers from a Bernoulli distribution  $\mathcal{B}(p)$ , where  $p$  is a hyperparameter controlling the proportion of  $\mathbf{x}$  to be modified. In both cases, the augmentation is done at the subject level.

**HemiMixUp:** This augmentation randomly permutes a subject’s measurements at specific vertices across hemispheres, assuming a vertex-to-vertex correspondence between hemispheres. Considering the left hemisphere, we get:

$$\tilde{\mathbf{x}}_i^{\text{left}} = \mathbf{m} \odot \mathbf{x}_i^{\text{right}} + (1 - \mathbf{m}) \odot \mathbf{x}_i^{\text{left}} \quad (2)$$

where  $\mathbf{x}_i^{\text{left}}$  and  $\mathbf{x}_i^{\text{right}}$  are the left and right hemisphere measures of the subject  $i$ , respectively.

**GroupMixUp:** The GroupMixUp augmentation randomly bootstraps measures at specific vertices across a group of  $K$  subjects  $G_i = \{g_1, \dots, g_K\}$  sharing similar cortical patterns with respect to the  $i$ -th subject. We aim to generate realistic noisy measures without missing hemispheric asymmetries by exploiting the group variability. We define  $\mathbf{X}_i = (\mathbf{x}_{g_1}, \dots, \mathbf{x}_{g_K})^T \in \mathbb{R}^{K \times P}$ . Considering the left hemisphere, we get:

$$\tilde{\mathbf{x}}_i^{\text{left}} = \mathbf{m} \odot \text{diag}[M \mathbf{X}_i^{\text{left}}] + (1 - \mathbf{m}) \odot \mathbf{x}_i^{\text{left}} \quad (3)$$

where  $\text{diag}$  is the diagonal operator, and  $M \in \{0, 1\}^{P \times K}$  is the random selection matrix. Each row of  $M$  selects a particular subject and is generated by drawing a random location from a uniform distribution  $\mathcal{U}(1, K)$ . The  $G_i$  grouping relies on a PCA trained on the residualized input data. The residualization is performed using ComBat to remove unwanted site-related noise [14]. We use  $K$ -nearest neighbors (with Euclidean distance) in the PCA space to define the group  $G_i$ . This step is performed only once before training, with little computational overhead. It is important to note that this strategy builds groups from a semantically meaningful space which maximizes the explained variance of the data. The first PCA axis is strongly related to age, as shown in Fig. 2-c. Groups are formed independently for each individual’s hemisphere.

By inverting the *left* and *right* notations in Eqs. 2 and 3, the formulations hold for the right hemisphere.

### 3 Experiments

#### 3.1 Data and Settings

**Datasets:** The T1 structural MRI data are processed with FreeSurfer, which calculates thickness, curvature, and sulcal morphology for each cortical vertex [13]. Interhemispheric registration (XHemi) is performed to obtain vertex-to-vertex mapping between hemispheres [15]. Inflated hemispheric cortical topologies are finally expressed on a regular order-5 icosahedral sphere. We use two datasets to demonstrate the proposed augmentations. First, we use the BHB, including more than 5000 individuals (age distribution  $25.3 \pm 15.0$ ) coming from multiple acquisition sites [11]. We use the so-called BHB internal train and test sets. We further split the BHB internal test set into validation and test sets (hereafter referred to as internal test), preserving the population statistics (age, sex, and acquisition site). Finally, we keep unchanged the BHB external set (hereafter referred to as external test), which consists of subjects with a similar age

distribution but disjoint acquisition sites. This set is used to evaluate generalization and robustness to unseen sites. Second, the Healthy Brain Network (HBN) cohort, which includes more than 1700 children (age distribution  $10.95 \pm 3.43$ ) with behavioral specificities or learning problems [1]. After applying the same quality control, we keep 1407 subjects. We split the data into training (80 %) and test (20 %) sets, preserving population statistics (age, sex and acquisition site). Some subjects (1073) have the cognitive score WISC-V FSIQ available.

**The Self-supervised Model:** SimCLR [5] contrastive learning strategy attempts to bring two representations of the same transformed sample as close as possible in the model latent space, while repelling other representations. We implement this model following the recent literature on self-supervised learning, which consists of an encoder and a projector. For the encoder, we choose a single SCNN architecture to facilitate the comparison between the methods (see Supplemental S3). It has four convolution blocks. Each convolution block consists of a DiNe convolution layer followed by a rectified linear unit and an average pooling operator [32]. There are two branches in the first convolution block, one for each hemisphere. The resulting features are concatenated on the channel axis and piped to the network flow. For the projector, we implement the architecture recommended in [5], a Multi-Layer Perceptron (MLP). The model is trained on the BHB training set. Three augmentation combinations are considered during training: all the baseline brain cortical augmentations (Base), Base + HemiMixUp, and Base + GroupMixUp. The entire procedure is repeated three times for each combination (nine trainings in total) to obtain a standard deviation for each prediction task described below.

**Model Selection and Evaluation:** In self-supervised learning, the training loss, even when evaluated on a validation set, indicates convergence but does not reflect the quality of the learned representations [22, 26]. As suggested in the literature to overcome this problem, we add a machine learning linear predictor on top of the encoder latent representations during training (ridge for regression and logistic for classification) (Fig. 1). We then estimate and monitor the associated prediction score from the validation set at each epoch (Mean Absolute Error (MAE) and coefficient of determination  $R^2$  for regression and Balanced Accuracy (BAcc) for classification). This score is only used to monitor the training, leaving the simCLR training process completely unsupervised. Finally, we evaluate the trained models with the same strategy for age and sex predictions on all cohorts and for FSIQ prediction on HBN. Age, sex and FSIQ are known to be proxy measures to investigate mental health [7]. They represent features that a pre-trained self-supervised model should be able to learn and generalize to unseen data. Due to the discrepancy in age distribution and the lack of clinical variables of interest, the linear predictors are fitted to the BHB and HBN training representations. Finally, the internal and external test sets and the HBN test set are used to evaluate the prediction scores.

**Table 1.** Evaluation of the learned representations using a machine learning linear predictor on different BHB (a)/HBN (b) sets of data, tasks, and metrics. The proposed MixUp augmentations (HemiMixUp and GroupMixUp) are evaluated against combined baseline (Base) augmentations (SurfCutOut, SurfBlur and SurfNoise) using an unsupervised SimCLR-SCNN framework. The results are compared to supervised SCNNs trained to predict either age or sex (see section **Model Comparison** for details).

a)

SimCLR-SCNN	BHB internal test			BHB external test		
Augmentations	Age		Sex	Age		Sex
	MAE ( $\downarrow$ )	$R^2$ ( $\uparrow$ )	BAcc ( $\uparrow$ )	MAE( $\downarrow$ )	$R^2$ ( $\uparrow$ )	BAcc ( $\uparrow$ )
Base	4.87 $\pm$ 0.14	0.81 $\pm$ 0.01	0.81 $\pm$ 0.01	5.89 $\pm$ 0.17	0.50 $\pm$ 0.03	0.71 $\pm$ 0.01
Base + HemiMixUp	4.72 $\pm$ 0.16	0.83 $\pm$ 0.01	0.81 $\pm$ 0.01	5.62 $\pm$ 0.13	0.55 $\pm$ 0.03	0.71 $\pm$ 0.02
Base + GroupMixUp	<b>4.55<math>\pm</math>0.07</b>	<b>0.84<math>\pm</math>0.01</b>	<b>0.82<math>\pm</math>0.01</b>	<b>5.47<math>\pm</math>0.15</b>	<b>0.58<math>\pm</math>0.02</b>	<b>0.74<math>\pm</math>0.01</b>
Supervised-SCNN						
Age-supervised	4.00 $\pm$ 0.12	0.84 $\pm$ 0.01	0.67 $\pm$ 0.01	5.06 $\pm$ 0.19	0.61 $\pm$ 0.02	0.55 $\pm$ 0.01
Sex-supervised	6.20 $\pm$ 0.20	0.69 $\pm$ 0.02	0.86 $\pm$ 0.01	6.40 $\pm$ 0.32	0.42 $\pm$ 0.06	0.68 $\pm$ 0.01

b)

SimCLR-SCNN	HBN test				
Augmentations	Age		Sex	FSIQ	
	MAE ( $\downarrow$ )	$R^2$ ( $\uparrow$ )	BAcc ( $\uparrow$ )	MAE( $\downarrow$ )	$R^2$ ( $\uparrow$ )
Base	1.69 $\pm$ 0.07	0.65 $\pm$ 0.02	0.80 $\pm$ 0.01	12.97 $\pm$ 0.25	0.10 $\pm$ 0.02
Base + HemiMixUp	1.68 $\pm$ 0.02	<b>0.66<math>\pm</math>0.01</b>	0.80 $\pm$ 0.01	<b>12.71<math>\pm</math>0.16</b>	<b>0.13<math>\pm</math>0.02</b>
Base + GroupMixUp	<b>1.66<math>\pm</math>0.02</b>	<b>0.66<math>\pm</math>0.01</b>	<b>0.81<math>\pm</math>0.01</b>	<b>12.60<math>\pm</math>0.29</b>	<b>0.14<math>\pm</math>0.03</b>
Supervised-SCNN					
Age-supervised	1.74 $\pm$ 0.05	0.63 $\pm$ 0.02	0.66 $\pm$ 0.01	13.05 $\pm$ 0.06	0.09 $\pm$ 0.002
Sex-supervised	1.72 $\pm$ 0.01	0.63 $\pm$ 0.01	0.82 $\pm$ 0.0048	12.79 $\pm$ 0.01	0.09 $\pm$ 0.01

**Model Comparison:** We compare the proposed SimCLR-SCNN model with supervised SCNNs. The supervised models consist of the same encoder followed by a linear predictor. The training loss for these supervised models depends on the task at hand (L1 for regression and cross-entropy for classification). These supervised models will be referred to as age-supervised if they were trained to predict age, and sex-supervised if they were trained to predict sex. Each supervised model is trained 3 times as well, to derive standard deviations, on the same train set as self-supervised models. They are evaluated the same way as self-supervised models: task-dependent machine learning linear predictors (ridge for regression and logistic for classification) are fitted to the learned representations from the SCNN encoder and evaluated on the test representations.

### 3.2 Results

**Self-supervised SCNNs Generalize Better Than Supervised SCNNs:** On BHB, compared to a much more specialized supervised SCNN setup, a



SCNN trained with the SimCLR self-supervised learning framework and the proposed augmentations shows a rather comparable performance for each of the investigated tasks (Table 1-a). For example, on the internal test, the SimCLR-SCNN age-MAE scores are 4.87, 4.72, and 4.55 for the Base, Base + HemiMixUp and Base + GroupMixUp augmentations, respectively. These scores are slightly worse than the predictions of the age-supervised SCNN (4.0 age-MAE), which is expected since the latter was trained in a supervised manner to learn good representations for predicting age. However they remain comparable and largely outperform the sex-supervised SCNN (6.2 age-MAE). The same trend can be observed for the  $R^2$  for the age prediction task and the BAcc for the sex prediction task for all test sets. Remarkably, for some tasks, the SimCLR-SCNN with the proposed augmentations even outperforms supervised SCNN models in terms of generalization performance. This can be seen by comparing the results on BHB internal and external test sets. For example SimCLR-SCNN loses 10% ( $0.81 \rightarrow 0.71$ ), 10% ( $0.81 \rightarrow 0.71$ ), and 8% ( $0.82 \rightarrow 0.74$ ) BAcc for the Base, Base + HemiMixUp and Base + GroupMixUp augmentations, between internal and external test sets, while age- and sex-supervised SCNNs lose 12% ( $0.67 \rightarrow 0.55$ ) and 18% ( $0.86 \rightarrow 0.68$ ) BAcc respectively. SimCLR SCNNs even outperform sex-supervised SCNNs for the sex prediction task on the external test set. As expected, the prediction on the external test decreases for both strategies. When the learned knowledge is transferred to HBN, we show better (or at least equivalent for sex) prediction performance for the SimCLR-SCNN with the proposed augmentations compared to the supervised SCNNs. Note that the age distribution in HBN is much narrower with a younger population than in BHB. Therefore, the MAE between Tables 1-a and 1-b cannot be directly compared. Interestingly, the  $R^2$  values are still comparable and show a stable goodness of fit for the SimCLR-SCNN model. Note that using a supervised MLP with more than 120M parameters to predict age from the same input data gives only slightly better results than using the SimCLR-SCNN ( $\sim 2$ M parameters) trained with the Base + GroupMixUp augmentations (4.85 *vs* 5.47 age-MAE on the BHB external test) [11]. This suggests that the SimCLR-SCNN model with the proposed augmentations is able to learn good representations without supervision and without being too much biased by the acquisition site.

**The MixUp Augmentations Improve Performance:** It is clear that the MixUp augmentations improve the learned representations for each prediction task of the BHB internal and external tests (Table 1-a). In practice, we found that the GroupMixUp works better than the HemiMixUp augmentation strategy. This can be explained by the attenuation of some properties of the inter-hemispheric asymmetry forced by the HemiMixUp augmentation. Although the improvement in predicting age and sex on HBN test set is inconclusive, it is clear that HemiMixUp and GroupMixUp help in predicting FISQ, especially when looking at the  $R^2$  metric (Table 1-b).

## 4 Conclusion

This study introduces cortical surface augmentations designed for training a SCNN in a self-supervised learning setup. The investigated SimCLR-SCNN shows the ability to generate representations with strong generalization properties. In fact, the learned representations from data collected from multiple sites offer promising performance, sometimes even outperforming supervised approaches. In particular, the GroupMixUp augmentation shows potential for learning stable representations across different cohorts. An ablation study could be used to further investigate the characteristics required for cortical-based data augmentation. Future work aims to incorporate prior information, such as clinical scores, into the GroupMixUp augmentation when computing the groups  $G_i$ . A similar strategy is proposed for structuring the learned representations by adding a regularization term in the training loss [10].

**Data Use Declaration and Acknowledgment.** The datasets analyzed during the current study are available online: OpenBHB in IEEEDataPort (doi 10.21227/7jsg-jx57), and HBN in NITRC (Release 10).

## References

1. Alexander, L.M., et al.: An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* **4**, 170181 (2017)
2. Alzubaidi, L., et al.: Towards a better understanding of transfer learning for medical imaging: a case study. *Appl. Sci.* **10**(13), 4523 (2020)
3. Bommasani, R., et al.: On the opportunities and risks of foundation models (2021)
4. Caron, M., et al.: Emerging properties in self-supervised vision transformers. In: *ICCV* (2021)
5. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. In: *ICML*, pp. 1597–1607 (2020)
6. Cohen, T.S., Geiger, M., Köhler, J., Welling, M.: Spherical CNNs. In: *ICLR* (2018)
7. Dadi, K., Varoquaux, G., Houenou, J., Bzdok, D., Thirion, B., Engemann, D.: Population modeling with machine learning can enhance measures of mental health. *GigaScience* **10**(10), giab071 (2021)
8. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. In: *NeurIPS*, vol. 29 (2016)
9. Defferrard, M., Milani, M., Gusset, F., Perraudin, N.: DeepSphere: a graph-based spherical CNN. In: *ICLR* (2020)
10. Dufumier, B., et al.: Contrastive learning with continuous proxy meta-data for 3D MRI classification. In: de Bruijne, M., et al. (eds.) *MICCAI 2021*. LNCS, vol. 12902, pp. 58–68. Springer, Cham (2021). [https://doi.org/10.1007/978-3-030-87196-3\\_6](https://doi.org/10.1007/978-3-030-87196-3_6)
11. Dufumier, B., Grigis, A., Victor, J., Ambroise, C., Frouin, V., Duchesnay, E.: OpenBHB: a large-scale multi-site brain MRI data-set for age prediction and debiasing. *Neuroimage* **263**, 119637 (2022)
12. Eitel, F., Schulz, M.A., Seiler, M., Walter, H., Ritter, K.: Promises and pitfalls of deep neural networks in neuroimaging-based psychiatric research. *Exp. Neurol.* **339**, 113608 (2021)

13. Fischl, B., Sereno, M.I., Dale, A.M.: Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* **9**(2), 195–207 (1999)
14. Fortin, J.P., et al.: Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* **161**, 149–170 (2017)
15. Greve, D.N., et al.: A surface-based analysis of language lateralization and cortical asymmetry. *J. Cogn. Neurosci.* **25**(9), 1477–1492 (2013)
16. Grill, J.B., et al.: Bootstrap your own latent - a new approach to self-supervised learning. In: *NeurIPS*, vol. 33 (2020)
17. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.: Momentum contrast for unsupervised visual representation learning. In: *CVPR* (2020)
18. Jiang, C.M., Huang, J., Kashinath, K., Prabhat, Marcus, P., Niessner, M.: Spherical CNNs on unstructured grids. In: *ICLR* (2019)
19. Khosla, P., et al.: Supervised contrastive learning. In: *NeurIPS* (2020)
20. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: *ICLR* (2017)
21. Raghu, M., Zhang, C., Kleinberg, J., Bengio, S.: Transfusion: understanding transfer learning for medical imaging. In: *NeurIPS*, vol. 32 (2019)
22. Saunshi, N., et al.: Understanding contrastive learning requires incorporating inductive biases. In: *ICML*, vol. 162, pp. 19250–19286 (2022)
23. Seong, S.B., Pae, C., Park, H.J.: Geometric convolutional neural network for analyzing surface-based neuroimaging data. *Front. Neuroinform.* **12**, 42 (2018)
24. Tian, Y., Sun, C., Poole, B., Krishnan, D., Schmid, C., Isola, P.: What makes for good views for contrastive learning? In: *Advances in Neural Information Processing Systems*, vol. 33, pp. 6827–6839. Curran Associates, Inc. (2020)
25. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lió, P., Bengio, Y.: Graph attention networks. In: *ICLR* (2018)
26. Wang, F., Liu, H.: Understanding the behaviour of contrastive loss. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, 19–25 June 2021*, pp. 2495–2504 (2021)
27. Yoon, J., Zhang, Y., Jordon, J., van der Schaar, M.: VIME: extending the success of self- and semi-supervised learning to tabular domain. In: *NeurIPS*, vol. 33 (2020)
28. Zbontar, J., Jing, L., Misra, I., LeCun, Y., Deny, S.: Barlow Twins: self-supervised learning via redundancy reduction. In: *ICML*, vol. 139 (2021)
29. Zhai, X., Oliver, A., Kolesnikov, A., Beyer, L.: S4L: self-supervised semi-supervised learning. In: *ICCV* (2019)
30. Zhang, H., Cissé, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: beyond empirical risk minimization. In: *6th International Conference on Learning Representations, ICLR 2018, Conference Track Proceedings, Vancouver, BC, Canada, 30 April–3 May 2018* (2018)
31. Zhao, F., et al.: Spherical deformable U-Net: application to cortical surface parcellation and development prediction. *IEEE Trans. Med. Imaging* **40**, 1217–1228 (2021)
32. Zhao, F., et al.: Spherical U-Net on cortical surfaces: methods and applications. In: *IPMI* (2019)