Prediction of Neuropsychological Scores from Functional Connectivity Matrices Using Deep Autoencoders

Delfina Irarte[1], Alberto Testolin[2](✉), Michele De Filippo De Grazia[2], and Marco Zorzi[2,3](✉)

[1] Department of Physics and Astronomy, University of Padova, 35141 Padua, Italy
[2] Department of General Psychology, University of Padova, 35141 Padua, Italy
{alberto.testolin,marco.zorzi}@unipd.it
[3] IRCCS San Camillo Hospital, 30126 Venice, Lido, Italy

Abstract. Deep learning models are being increasingly used in precision medicine thanks to their ability to provide accurate predictions of clinical outcome from large-scale datasets of patient's records. However, in many cases data scarcity has forced the adoption of simpler (linear) feature extraction methods, which are less prone to overfitting. In this work, we exploit data augmentation and transfer learning techniques to show that deep, non-linear autoencoders can in fact extract relevant features from resting state functional connectivity matrices of stroke patients, even when the available data is modest. The latent representations extracted by the autoencoders can then be given as input to regularized regression methods to predict neurophsychological scores, significantly outperforming recently proposed methods based on linear feature extraction.

Keywords: Resting state networks · Functional connectivity · Deep learning · Feature extraction · Predictive modeling

1 Introduction

Improvements in neuroimaging have provided physicians and radiologists with the ability to study the brain with unprecedented precision. In particular, Resting State functional Magnetic Resonance Imaging (RS-fMRI) measures spontaneous fluctuations in blood oxygen-level dependent neural activity and allows estimating the brain functional connectivity in the absence of any task-related activity [1].

Functional connectivity of resting state networks has shown to be a valuable predictor of individual neuropsychological scores in stroke survivors, making it a potentially useful tool in clinical practice [2–4]. However, building robust predictive models from such high-dimensional measurements requires a large number of training samples, which are not always available in clinical populations. Such limitation can be partially addressed by exploiting linear dimensionality reduction techniques such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), or sparse coding in combination with regularized regression

M. Mahmud et al. (Eds.): BI 2022, LNAI 13406, pp. 140–151, 2022
https://doi.org/10.1007/978-3-031-15037-1_12

methods [5,6]. Nevertheless, the choice of the dimensionality reduction technique is non-trivial because it can affect performance of the predictive model [5,7].

Here we show that better performance can be achieved by exploiting the representational power of non-linear dimensionality reduction techniques, namely, deep autoencoders [8]. Autoencoders (AE) are becoming popular in functional neuroimaging thanks to their ability to disentangle the underlying brain dynamics in a completely unsupervised way [9,10] and have already been successfully used to build predictive models of psychiatric disorders [11,12]. Nevertheless, the application of such powerful deep learning models is often hindered by the limited size of clinical datasets. In this work we propose to mitigate this issue using two complementary approaches: data augmentation, which allows to expand the sample size by combining/distorting existing samples, and transfer learning, which allows to exploit additional large-scale datasets (in our case, from the Human Connectome Project [13]) containing functional connectivity data in order to pre-train the autoencoder.

The proposed approach is validated on a reference dataset containing functional connectivity matrices of stroke patients [3]. The features extracted by the autoencoder are used as predictors of the corresponding neurophsychological scores by means of regularized linear regression methods. The latter can limit multicollinearity and overfitting, which makes them particularly suitable for the analysis of neuroimaging data (for a recent review, see [14]). The performance of our method is benchmarked against other popular dimensionality reduction methods based on PCA and ICA, showing promising results.

2 Materials and Methods

2.1 Datasets

The main dataset used in our study consists of 100 resting state functional connectivity (RSFC) matrices from symptomatic stroke patients, taken from previous studies [3,5]. The patients underwent a 30-minute-long RS-fMRI acquisition, 1–2 weeks after the stroke occurred. Several scores were taken during the neuropsychological assessment: here we focus on language, verbal memory and spatial memory indexes, which are available for a subset of subjects (language: N = 94; memory: N = 77). In order to implement a transfer learning approach, we also used a dataset from the Human Connectome Project [13], consisting of RSFC matrices of 1050 healthy subjects. RSFC data represent the connectivity between brain regions that share functional properties and can be expressed as a symmetric matrix. In our case, the matrix of each subject is of size 324×324; following common practice [5], the data was vectorized by only considering the upper triangular matrix. Null values were converted to zero.

2.2 Dimensionality Reduction

Dimensionality reduction is the process of taking some input data in a high dimensional space and mapping it into a new "feature" space whose dimensionality is much smaller [15]. Our main focus was to test different variants of deep autoencoders in their ability to extract useful features from RSFC data, and compare their performance with standard linear dimensionality reduction methods [5]. The models were initially compared in terms of their reconstruction error, which corresponds to the mean squared error between the original matrix and the reconstructed one. During the unsupervised feature extraction process, the entire dataset (n = 100) was used regardless of the availability of neuropsychological scores.

Linear dimensionality reduction techniques, such as PCA and ICA, apply a linear transformation to the input data. That is, if the original data is in $\mathcal{R}^d$ and we want to embed it into $\mathcal{R}^n$ $(n < d)$ then we would like to find a matrix $W \in \mathcal{R}^{n,d}$ that induces the mapping $x \rightarrow Wx$. A natural criterion for choosing W is in a way that will enable a reasonable recovery of the original input x [15]. Compared to deep autoencoders, the main drawback of PCA and ICA is that they cannot extract nonlinear structures modeled by higher than second-order statistics [16]. In the following, we will briefly review the main techniques used in the present study and their implementation.

Principal Component Analysis. Before performing PCA the data was standardized to obtain a distribution with zero mean and unit variance. This step was implemented using the predefined function `StandardScaler` from SKLEARN. PCA was then performed by using the function `PCA` from the same library, which performs linear dimensionality reduction using Singular Value Decomposition of the data to project it to a lower dimensional space.

Independent Component Analysis. ICA performs the decomposition step by imposing the constraint that the resulting components must be independent. In this work we used the `FastICA` algorithm from SKLEARN, which is a block fixed-point iteration algorithm based on negative entropy as a non-gaussianity measure, which converges faster than adaptive algorithms [9]. As in the case of PCA, data was first standardized.

Autoencoders. An autoencoder is an unsupervised neural-network based approach for learning latent representations of high-dimensional data that can be used to reconstruct the original input, while compressing it into a latent space that usually has much lower dimensionality [17]. Learning such "undercomplete" representations forces the autoencoder to capture the most salient features of the training data by discovering its latent factors of variation [18].

Let's consider a basic auto-encoder with a single hidden layer, n neurons in the input/output layers and m neurons in the hidden layer. The model takes an input $\mathbf{x} \in \mathcal{R}^n$ and first maps it into the latent representation $\mathbf{h} \in \mathcal{R}^m$ by using an encoding function $\mathbf{h} = g_\phi(\mathbf{x}) = \sigma(W\mathbf{x}+b)$ with parameters $\phi = \{W, b\}$, where $\sigma(\cdot)$ denotes the activation function of the neurons, W denotes the connection weights and b denotes the neurons' biases. Afterwards, a reconstruction of the input $\mathbf{x}'$ is obtained through the decoder function $\mathbf{x}' = f_\theta(\mathbf{h}) = \sigma(W'\mathbf{h} + b')$

with $\theta = \{W', b'\}$. The two parameter sets (θ, ϕ) are usually constrained to be of the form $W \in \mathcal{R}^{n,m} = W'^{T} \in \mathcal{R}^{m,n}$, using the same weights for encoding the input and decoding the latent representation [19]. The parameters are learned by minimizing an appropriate cost function over the training set, which usually corresponds to the Mean Squared Error between the original input and the reconstructed output:

$$L_{\mathrm{AE}}(\theta, \phi) = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{x}^{(i)} - f_{\theta}(g_{\phi}(\mathbf{x}^{(i)})))^{2} \tag{1}$$

Fully connected AE do not have any spatial bias over the image structure. Convolutional autoencoders are an AE variant that exploits convolution filters to more efficiently capture local spatial structure. For a mono-channel input x the latent representation of the k$-th$ feature map is given by:

$$h^{k} = \sigma(x * W^{k} + b^{k}) \tag{2}$$

where the bias is broadcasted to the whole feature map, σ is an activation function, and $*$ denotes a convolution. The reconstruction is obtained using:

$$y = \sigma(\sum_{k \in H} h^{k} * \hat{W}^{k} + c) \tag{3}$$

where c represents the bias of the input channel, H identifies the group of latent feature maps and $\hat{W}$ identifies the flip operation over both dimensions of the weights [19].

In this work we considered both fully-connected and Convolutional Autoencoder (CAE) architectures. As baselines, we implemented two simple, 1-layer AE with linear and non-linear activation functions. We then implemented a more sophisticated CAE architecture, as shown in Fig. 1. In the latter case, the encoder consisted of 3 convolutional layers followed by 2 fully connected layers, and the same structure was mirrored in the decoder. In order to overcome vanishing gradient the Leaky Rectified Linear activation function was used. Mean Square Error was used as loss function, which was minimized using the Adam optimizer with a learning rate of $1e{-}3$. Dropout was used as a further regularizer. Hyperparameters were automatically optimized using Optuna [20].

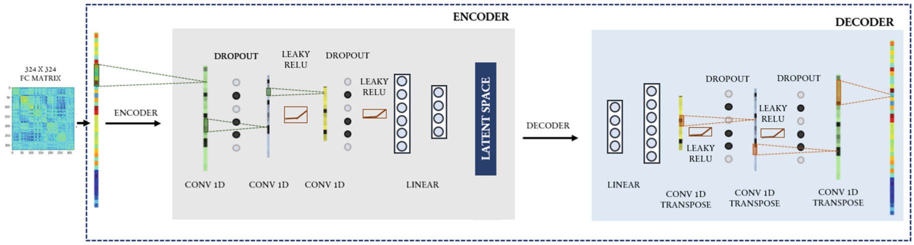

Fig. 1. Workflow and architecture of the deep convolutional autoencoder.

2.3 Data Augmentation and Transfer Learning

Deep networks perform remarkably well in many domains, but they are heavily reliant on big data to avoid overfitting. Given the limited size of our clinical dataset, we thus devised two approaches in order to promote a better generalization of the CAE during the feature extraction process.

The first method was based on *Data Augmentation*, which consists in combining and distorting each training sample in order to provide a more representative distribution as input to the autoencoder [21]. In particular, we designed a mix-up augmentation method that consists of a random convex combination of two input samples leading to a total of 7421 synthetic samples:

$$\hat{x} = \lambda x_i + (1 - \lambda) x_j$$

where x_i and x_j are raw input vectors and λ are values sampled from the Beta distribution[1]. Following previous work [22], the choice of the parameters $\lambda \in [0, 1]$ was distributed accordingly to $\lambda \in Beta(\alpha, \alpha)$ for $\alpha \in (0, \inf)$. In the mix-up, the samples to be combined were chosen randomly from all available images. Isaksson et al. [23] tested the utility of the mix-up data augmentation technique for a medical image segmentation task using 100 MRI scans and observed an improvement when $\alpha = 0.5$. Although our dataset could be slightly different, we decided to use the same α value for consistency.

The second method was based on *Transfer Learning (TL)*, which consists in first training the autoencoder on a larger-scale dataset and subsequently tune it on the smaller dataset. In our case, we took advantage of the Human Connectome Project database for the pretraining phase. Afterwards, the model was fine-tuned using the stroke dataset freezing the weights of the convolutional layers.

2.4 Regularized Regression

The feature sets extracted by each method were used as regressors for the prediction of the neuropsychological scores.

Ridge regression [24] is a regularized regression method that controls the regression coefficients by adding the L_2 penalty term $\lambda \sum_{j=1}^{p} \beta_j^2$ to the objective function. The least absolute shrinkage and selection operator (LASSO) model [25] is an alternative method that adds the L_1 penalty term $\lambda \sum_{j=1}^{p} |\beta_j|$. To implement regularized regression we exploited a flexible approach based on elastic net [26], which combines the penalties of Ridge and LASSO regression:

$$\min_{(\beta_0, \boldsymbol{\beta})} \left(\boldsymbol{y} - \beta_0 - \boldsymbol{X}^T \boldsymbol{\beta} \right)^2 + \lambda \left(\frac{1}{2} (1 - \alpha) \boldsymbol{\beta}^2 + \alpha |\boldsymbol{\beta}| \right), \tag{4}$$

The elastic-net loss function requires two free parameters to be set, namely λ and α. The penalty parameter λ controls the amount of shrinkage, while the parameter α controls the type of shrinkage. Following previous work [5], these

[1] Note that although the extracted features were obtained using the synthetic data, the model performance was always measured on the final stroke dataset.

parameters were tuned using Leave-One-Out Cross validation (LOOCV). To evaluate the regression model we used R-squared (R^2), Mean Squared Error (MSE) and Bayesian information criterion (BIC).

3 Results

3.1 Dimensionality Reduction

Figure 2 shows the reconstruction error against the number of components/latent units for each method. The trend is similar across models: the larger the number of components, the better the reconstruction. The CAE trained directly on the stroke dataset obtained the worst reconstruction error, while the CAE trained on the augmented dataset achieved the best performance. This result highlights the importance of increasing the variability of the training distribution in order to improve the quality of the features extracted by complex convolutional architectures. The simple 1-layer AEs achieved an intermediate reconstruction error, comparable to those of PCA and ICA, which is no surprise given the intrinsic similarity between these techniques [27].

3.2 Regularized Regression

Table 1 presents the metrics obtained in the neuropsychological scores prediction task. As it can be observed, the λ parameter is usually small. On the other hand, it can be seen that the α value mainly takes the two extremes: $\alpha \sim 1$, which corresponds to a ridge regression; and $\alpha \sim 0$, which corresponds to a LASSO penalization; an intermediate $\alpha \sim 0.75$ only happens in few cases. In order to have a better visualization, Fig. 3 presents the methods sorted by lowest MSE error and highest R^2.

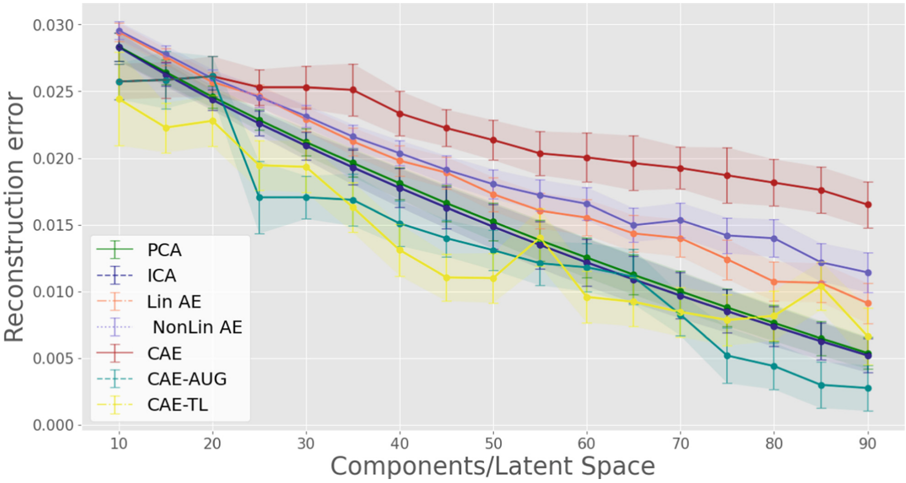

Fig. 2. Reconstruction error achieved by different feature extraction methods.

Concerning the metrics obtained for the language score, it can be observed that PCA slightly outperforms the other models in terms of R^2 and MSE, though the margin is fairly small. However, the CAE trained with Data Augmentation achieves the best performance in both spatial and memory scores, with a considerable margin over the other methods. Such remarkable performance is approached also by the CAE trained using Transfer Learning. Interestingly, the autoencoder with a single linear layer is often the one achieving the lowest BIC value, suggesting that such architecture is particularly useful to select a few representative components from the data distribution.

3.3 Getting Deeper on Augmentation and Transfer Techniques

Given the remarkable performance of the CAE trained using data augmentation and transfer learning, in a series of additional simulations we explored how the size of the augmented dataset could impact model performance, and whether a combination of data augmentation and transfer learning might further improve the predictive accuracy[2]. We thus designed four additional training regimens:

Table 1. Regression metrics and parameters obtained for the different feature extraction methods.

	Language score (n = 94)					Spatial score (n = 77)					Memory score (n = 77)				
	R^2	MSE	BIC	α	λ	R^2	MSE	BIC	α	λ	R^2	MSE	BIC	α	λ
PCA	**0.52**	**0.48**	493	0.00	0.22	0.21	0.79	**300**	1	0.09	0.32	0.68	363	1	0.03
ICA	0.51	0.49	351	0.25	0.09	0.24	0.75	396	0.00	0.56	0.27	0.73	381	1	0.04
Lin AE	0.43	0.57	**323**	0.25	0.06	0.27	0.73	412	0.00	0.56	0.25	0.75	**297**	0.5	0.15
NonLin AE	0.50	0.50	357	0.75	0.00	0.26	0.74	456	0.00	0.22	0.26	0.74	369	0.75	0.01
CAE	0.42	0.57	624	0.25	0.01	0.27	0.73	390	0.5	0.01	0.27	0.73	759	0.00	0.7
CAE-AUG	0.50	0.50	421	0.50	0.06	**0.33**	**0.65**	315	0.5	0.09	**0.40**	**0.61**	316	1	0.04
CAE-TL	0.44	0.56	454	0.00	0.03	0.31	0.69	407	0.75	0.00	0.39	0.61	302	0.75	0.01

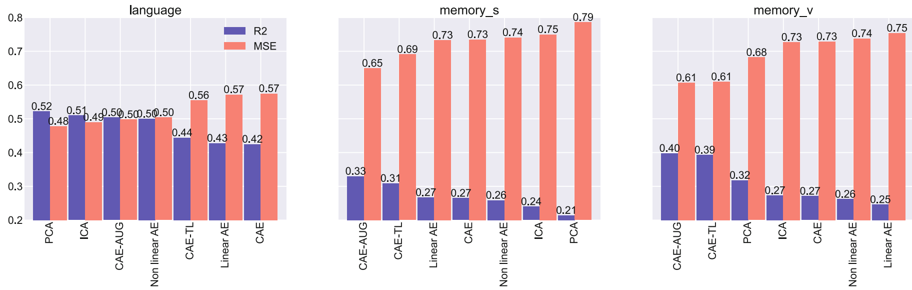

Fig. 3. MSE (orange) and R^2 (violet) metrics obtained by different methods sorted by accuracy. (Color figure online)

[2] It should be pointed out that for these simulations we did not implement an exhaustive hyper-parameter optimization, as in the previous cases.

1. **Aug(15000)**: Similarly as before, the CAE is trained with synthetic images obtained via the mix-up strategy; however this time the size of the augmented stroke dataset is increased to ∼15000 samples (i.e., twice the size used previously);
2. **TL-Aug**: The CAE is first trained over the HCP dataset, as done before for the Transfer Learning scenario. The model is then also trained on the initial augmented stroke dataset (∼7500 samples);
3. **AugTL-Aug:** The CAE is first trained over synthetic HCP data obtained by applying the same mix-up augmentation strategy (∼6000 samples). The model is then also trained on the initial augmented stroke data (∼7500 samples);
4. **AugTL-Stroke:** The CAE is first trained over synthetic HCP data obtained by applying the same mix-up augmentation strategy (∼6000 samples). The model is then also trained on the original stroke dataset.

Figure 4 shows the reconstruction error obtained by the four different regimens. The errors are comparable to that achieved previously by the simpler Data Augmentation technique, suggesting that also in these cases we achieve very good reconstructions.

At the same time, regression results reported in Table 2 and Fig. 5 clearly show that these improved data augmentation and transfer learning regimens further boosted the model's performance, both in terms of R^2 and MSE. All regimens generally enhance the CAE accuracy, however the most striking improvement is given by the TL-Aug regimen, which reaches significantly better performance compared to all methods previously investigated, establishing a new state-of-the-art for the stroke-prediction task. Interestingly, this improved model achieves such accurate predictions by relying, on average, on fewer components compared to other methods, which might be particularly relevant to improve interpretability of the resulting model.

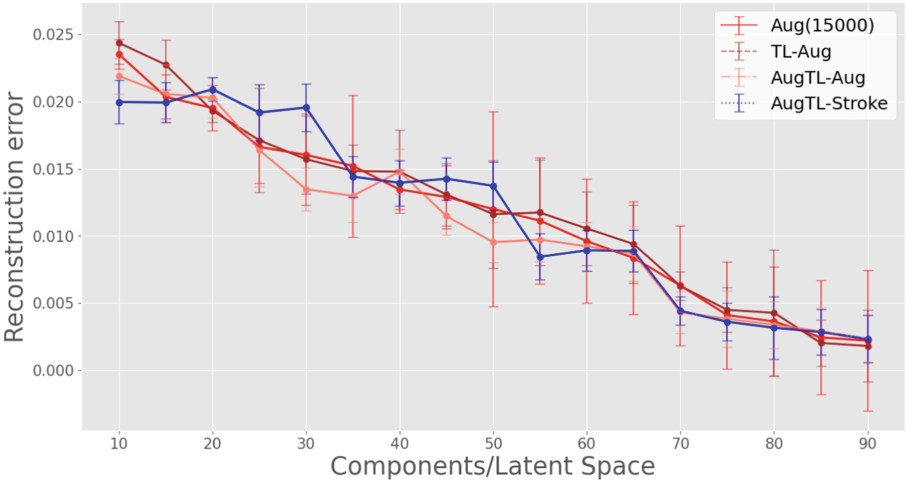

Fig. 4. Reconstruction error achieved by the four new augmentation/transfer regimens.

Table 2. Regression metrics and parameters obtained by the four augmentation/transfer regimens.

	Language score (n = 94)					Spatial score (n = 77)					Memory score (n = 77)				
	R^2	MSE	BIC	α	λ	R^2	MSE	BIC	α	λ	R^2	MSE	BIC	α	λ
Aug (15000)	0.51	0.49	421	0.5	0.06	0.36	0.58	570	0.00	0.05	0.41	0.59	570	0.00	0.05
TL-Aug	**0.56**	**0.45**	284	0.00	0.03	**0.40**	**0.56**	367	0.5	0.09	**0.47**	**0.54**	357	0.75	0.08
AugTL-Aug	0.53	0.46	421	0.5	0.06	0.23	0.77	247	1	0.16	0.43	0.57	239	1	0.08
AugTL-Stroke	0.47	0.53	433	1	0.02	0.28	0.72	380	0.00	0.81	0.42	0.58	242	1	0.16

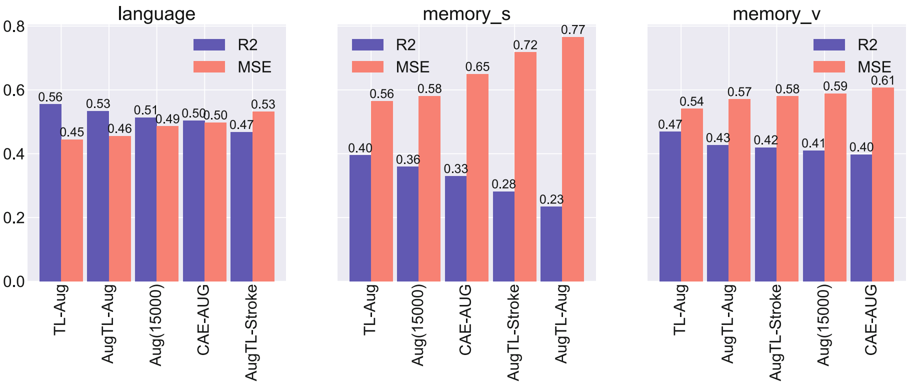

Fig. 5. MSE and R^2 metrics obtained by augmentation/transfer regimens sorted by accuracy.

4 Conclusion

In this work we investigated whether deep autoencoders could extract relevant features from resting state functional connectivity data of stroke patients, which can successively be used to build predictive models of neuropsychological scores. We implemented a variety of autoencoder architectures, ranging from simple, one-layer linear networks to more sophisticated convolutional versions exploiting several layers of non-linear processing. In order to deal with the issue of data scarcity, which is known to affect the performance of deep learning models, we also explored data augmentation and transfer learning techniques. The autoencoder's performance was benchmarked against other conventional approaches, such as Principal Component Analysis (PCA) and Independent Component Analysis (ICA).

The different methods were first evaluated in terms of their reconstruction error. In general, all methods achieved similar reconstruction error, though the autoencoders trained using data augmentation obtained slightly better accuracy. The quality of the features extracted by different methods was then assessed based on their capacity to serve as predictors for neuropsychological scores of the patients in three cognitive domains (i.e., language, spatial memory, and verbal memory). To this aim, the extracted features were given as input to regularized regression models, and performance was evaluated in terms of coefficient of determination, mean squared error and Bayesian information criterion. Results showed that the performance of the basic autoencoders was overall comparable to that of traditional methods (ICA and PCA). However, more sophisticated convolutional architectures trained using data augmentation and transfer learning achieved a much higher performance, with considerable gains of 7% (language), 66% (spatial memory) and 47% (verbal memory) with respect to the previously reported state-of-the-art methods [5]. The larger accuracy gains for memory scores can be explained by the fact that prediction of language scores is likely close to ceiling. Memory has a more distributed neural basis and the prediction of deficits from structural lesions is relatively poor compared to other behavioral domains [4,6]. Therefore, predicting memory scores represents an important benchmark for RSFC-based machine learning methods.

In conclusion, our results demonstrate the great potential of deep learning models for the analysis of multi-dimensional neuroimaging data even in cases with limited data availability, which is often considered a critical limitation in clinical studies. Future work should aim at further consolidating our findings, for example by systematically evaluating the performance of deep learning models on the prediction of other neuropsychological and behavioral scores, or by increasing the sample size in order to allow testing model generalization on fully held-out data. The latter task calls for multi-centric, coordinated efforts for collection, harmonization and sharing of patients' functional imaging data. Moreover, a key research frontier would be to design and implement advanced techniques in order to interpret the features extracted by non-linear "black-box" models, such as deep networks. Although standard back-projection techniques [5] only work with linear dimensionality reduction, there is a growing interest in designing explainability techniques that can visualize the features that mostly influence the decision of deep networks (for a recent review, see [28]). Such techniques would be particularly relevant in the case of medical applications, since they could provide valuable insights to the clinicians for the design of more effective rehabilitation protocols.

Acknowledgements. This work was supported by grants from the Italian Ministry of Health (RF-2019-02359306 to MZ, Ricerca Corrente to IRCCS Ospedale San Camillo) and by MIUR (Dipartimenti di Eccellenza DM 11/05/2017 n. 262 to the Department of General Psychology). We are grateful to Prof. Maurizio Corbetta for providing the stroke dataset, which was collected in a study funded by grants R01 HD061117-05 and R01 NS095741. Healthy adults rs-fMRI data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

1. Greicius, M., Supekar, K., Menon, V., Dougherty, R.: Resting-state functional connectivity reflects structural connectivity in the default mode network. Cereb. Cortex **19**, 72–8 (2008)
2. Meskaldji, D.E., et al.: Prediction of long-term memory scores in MCI based on resting-state FMRI. NeuroImage Clin. **12**, 785–795 (2016)
3. Siegel, J.S., et al.: Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. Proc. Natl. Acad. Sci. **113**(30), E4367–E4376 (2016)
4. Salvalaggio, A., De Filippo De Grazia, M., Zorzi, M., Thiebaut de Schotten, M., Corbetta, M.: Post-stroke deficit prediction from lesion and indirect structural and functional disconnection. Brain **143**(7), 2173–2188 (2020)
5. Calesella, F., Testolin, A., De Filippo De Grazia, M., Zorzi, M.: A comparison of feature extraction methods for prediction of neuropsychological scores from functional connectivity data of stroke patients. Brain Inform. **8**, 1–13 (2021)
6. Zorzi, M., De Filippo De Grazia, M., Blini, E., Testolin, A.: Assessment of machine learning pipelines for prediction of behavioral deficits from brain disconnectomes. In: Mahmud, M., Kaiser, M.S., Vassanelli, S., Dai, Q., Zhong, N. (eds.) BI 2021. LNCS (LNAI), vol. 12960, pp. 211–222. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86993-9_20
7. Jollans, L., et al.: Quantifying performance of machine learning methods for neuroimaging data. Neuroimage **199**, 351–365 (2019)
8. Bank, D., Koenigstein, N., Giryes, R.: Autoencoders. arXiv abs/2003.05991 (2020)
9. Kim, J.H., Zhang, Y., Han, K., Wen, Z., Choi, M., Liu, Z.: Representation learning of resting state fMRI with variational autoencoder. NeuroImage **241**, 118423 (2021)
10. Huang, H., et al.: Modeling task fMRI data via deep convolutional autoencoder. IEEE Trans. Med. Imaging **37**(7), 1551–1561 (2017)
11. Pinaya, W., Mechelli, A., Sato, J.: Using deep autoencoders to identify abnormal brain structural patterns in neuropsychiatric disorders: a large-scale multi-sample study. Hum. Brain Mapp. **40**, 944–954 (2018)
12. GENG, X.F., Xu, J.: Application of autoencoder in depression diagnosis. DEStech Trans. Comput. Sci. Eng. (2017)
13. Van Essen, D.C., et al.: The WU-Minn human connectome project: an overview. Neuroimage **80**, 62–79 (2013)
14. Cui, Z., Gong, G.: The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. Neuroimage **178**, 622–637 (2018)
15. Shalev-Shwartz, S., Ben-David, S.: Understanding Machine Learning - From Theory to Algorithms. Cambridge University Press, Cambridge (2014)
16. Cai, B., et al.: Functional connectome fingerprinting: identifying individuals and predicting cognitive functions via autoencoder. Hum. Brain Mapp. **42**, 2691–2705 (2021)
17. Pedrycz, W., Chen, S.-M. (eds.): Deep Learning: Algorithms and Applications. SCI, vol. 865. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-31760-7
18. Scholz, M., Vigário, R.: Nonlinear PCA: a new hierarchical approach. In: ESANN (2002)
19. Masci, J., Meier, U., Cireşan, D., Schmidhuber, J.: Stacked convolutional autoencoders for hierarchical feature extraction. In: Honkela, T., Duch, W., Girolami, M., Kaski, S. (eds.) ICANN 2011. LNCS, vol. 6791, pp. 52–59. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21735-7_7

20. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2019)
21. Shorten, C., Khoshgoftaar, T.: A survey on image data augmentation for deep learning. J. Big Data **6**, 1–48 (2019)
22. Zhang, H., Cissé, M., Dauphin, Y., Lopez-Paz, D.: Mixup: beyond empirical risk minimization. arXiv abs/1710.09412 (2018)
23. Isaksson, L., et al.: Mixup (sample pairing) can improve the performance of deep segmentation networks. J. Artif. Intell. Soft Comput. Res. **12**, 29–39 (2022)
24. Hoerl, A., Kennard, R.: Ridge regression: biased estimation for nonorthogonal problems. Technometrics **12**, 55–67 (2012)
25. Tibshirani, R.: Regression shrinkage selection via the lasso. J. Roy. Stat. Soc. Ser. B **73**, 273–282 (2011)
26. Zou, H., Hastie, T.: regularization and variable selection via the elastic net. J. Roy. Stat. Soc. B **67**(2), 301–320 (2005)
27. Baldi, P., Hornik, K.: Neural networks and principal component analysis: learning from examples without local minima. Neural Netw. **2**(1), 53–58 (1989)
28. Singh, A., Sengupta, S., Lakshminarayanan, V.: Explainable deep learning models in medical image analysis. J. Imaging **6**(6), 52 (2020)