

# Toward more frugal models for functional cerebral networks automatic recognition with resting state fMRI

Lukman E. ISMAILA<sup>1</sup>, Pejman RASTI<sup>1,2</sup>, Jean-Michel LEMÉE<sup>3</sup>, David ROUSSEAU<sup>1\*</sup>

<sup>1</sup>LARIS, UMR INRAe, IRHS Angers, Université d'Angers, France

<sup>2</sup>CERADE ESAIP, École d'Ingénieurs

<sup>3</sup>Service de Neurochirurgie CHU d'Angers, France

david.rousseau@univ-angers.fr

**Résumé** – Nous considérons une situation d'apprentissage machine où des modèles à base de réseaux de neurones convolutionnels classiques ont montré de bonnes performances. Nous investiguons différentes techniques d'encodage sous forme de supervoxel puis de graphes pour réduire la complexité du modèle tout en limitant la perte de performance. Cette approche est illustrée sur une tâche de reconnaissance de réseaux fonctionnels de repos pour des patients atteints de tumeurs cérébrales. Les graphes encodant des supervoxels préservent les caractéristiques d'activation des réseaux cérébraux fonctionnels à partir des images, réduisant les paramètres du modèle de 26 fois tout en maintenant les performances du modèle CNN.

**Abstract** – We refer to a machine learning situation where models based on classical convolutional neural networks have shown good performance. We are investigating different encoding techniques in the form of supervoxels, then graphs to reduce the complexity of the model while tracking the loss of performance. This approach is illustrated on a recognition task of resting-state functional networks for patients with brain tumors. Graphs encoding supervoxels preserve activation characteristics of functional brain networks from images, optimize model parameters by 26 times while maintaining CNN model performance.

## 1 Introduction

Convolutional neural networks (CNN) are powerful tools to perform computer vision tasks. CNN are however very demanding in terms of energy, data and annotation due to the large amount of parameters to be tuned during their training. These limitations are specially important in medical imaging where the constitution of large cohorts of unhealthy patients can be a bottleneck as frequently observed in cases of rare diseases like brain tumor. Recently, we have shown the possibility to circumvent this limitation by the use of transfer learning from self-supervised training on healthy data to unhealthy data [1]. We used small data in our experiments, and approach opens the possibility for scalability when a larger model is trained from additional data acquired.

This was obtained for the automatic recognition of functional cerebral networks via resting-state functional magnetic resonance imaging (rs-fMRI) [2] for patient with brain tumors. The CNN architecture proposed for the classification of functional brain network with 3D fMRI images by Ismaila *et al.*, was observed with high model training parameters despite the small data size [2] which constitutes a complex model and struggles with risks of overfitting.

In this work, we test possible ways to simplify deep learning models by reducing the overall parameter size. To this purpose, we propose to compare a basic CNN method with the

approach depicted in Fig. 1. Based on a recent work by Gousia *et al.*, which highlighted the benefits of graph encoding in optimizing CNN model parameters especially in medical imaging [3]. We investigate various ways of encoding the rs-fMRI 3D volume data in more compacted fashions and systematically compare our observation with the performance obtained in [2]. This effort only represent an initial attempt towards more efficient encoding of our brain volume images, as well as opens the possibility for scalability when a larger model is trained from additional data acquired.

## 2 Database

fMRI brain network activation image data of 81 healthy subjects and 55 unhealthy patients were collected. Regular volunteers provide the healthy data, while patients with brain tumors where a binary mask indicate region of lesion in the brain constitute the unhealthy data. This analysis, was done in separate components which creates brain maps of the regions with synchronous blood oxygen level dependent (BOLD) signal activity. In the data acquisition stage, we extracted the intrinsic connectivity networks (ICNs) by using methods that combine the information of both the temporal and spatial dimensions, such as independent component analysis. The extracted signals represent the neuro-anatomical basis for the functional

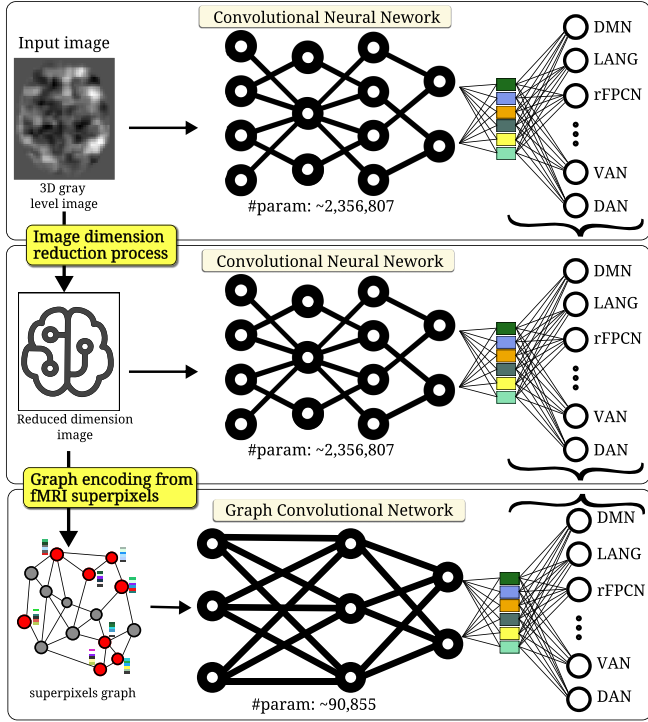


Figure 1: Visual abstract of our method. We consider as baseline performance either in terms of number of parameters and accuracy of CNN applied on raw rs-fMRI activation maps for functional cerebral network automatic recognition. We compare this performance with neural networks applied on compacted versions of the images.

networks in the brain [4]. The statistical parametric mapping (SPM) anatomy toolbox for Matlab was used to generate the 3D brain volume images, from the initial spatio-temporal fMRI signals. Among the 55 ICNs processed for each patients, 7 of these signals were recognized manually by experts to be biological networks of the brain such as Default Mode Network (DMN), Language Network (LANG), Right Frontoparietal Control Network (rFPCN), Left Frontoparietal Control (lFPCN), Salience Network (SAL), Dorsal Attention Network (DAN) and Ventral Attention Network (VAN). The annotated images were used in two versions: full images (connectivity map) and corresponding thresholded images.

### 3 Spatial dimension reduction

One may wonder if the entire 3D volume in gray levels is fully informative for automatic recognition of the functional cerebral networks. Several dimension reduction approaches can be envisioned. From the acquired brain volumes of resting-state fMRI images  $42px \times 51px \times 34channels$ , we normalized the pixel intensity range to 0-1 and computed several reduced version of these raw data as depicted in Fig. 2.

First, one can reduce the number of spatial dimension via a projection. We produced 2D gray level image by performing

Mean operation on pixel intensity across the axial (A) plane as shown in Fig. 2 Secondly, to understand whether the intensity of the activation map holds discriminative information, we created 2D binary images by performing an OR operation in respect to sagittal, coronal and axial (SCA) plane respectively, which were further stacked together to provide SCA binary stack image. Also, we performed another OR operation across the axial plane to obtain a 3D binary volume image which overall, resulted in 4 variants of generated images as illustrated in Fig. 2. Lastly, we tested if the full resolution of voxels is necessary for the classification of the functional network, which are rather formed by large structures than fine details. To this purpose, segmentation of the gray level activation map was performed using SLIC algorithm [5, 6]. We processed the 2D segmented labels to obtain a superpixels image, while the 3D segmented labels provided the supervoxels image as shown in Fig. 4. Furthermore, we averaged (smoothened) the pixel intensities within each segment of our superpixels and supervoxels images. This step allows us to evaluate the integrity of the functional brain network features which was done by training a CNN model for 7 distinct functional brain network classification using the generated superpixels/supervoxels images.

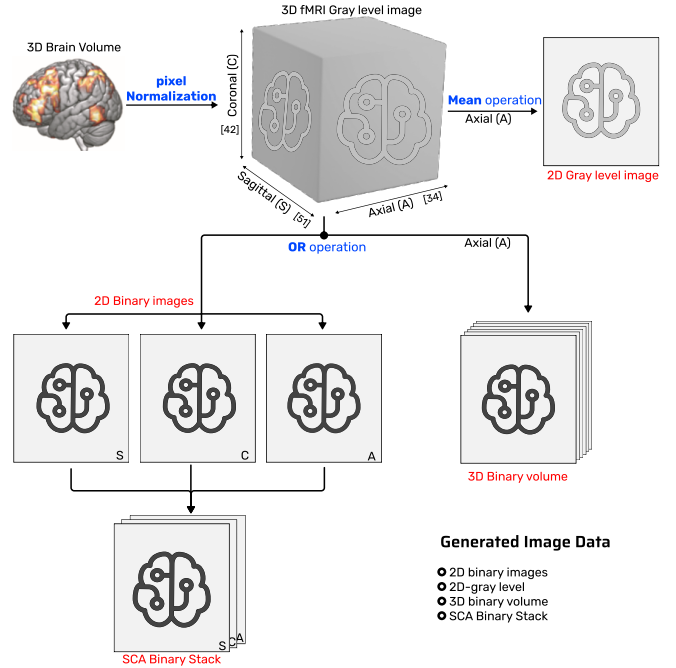


Figure 2: fMRI image dimension reduction process.

Table 1: CNN based fMRI brain network classification with unhealthy data.

Data	Train-Test	Accuracy	Parameters
3D gray level	315-70	0.75 ± 0.01	2,356,807
3D binary	315-70	0.66 ± 0.02	2,356,807
2D gray level	315-70	0.68 ± 0.01	2,337,799
2D binary	315-70	0.63 ± 0.01	2,337,799
SCA-binary stack	315-70	0.68 ± 0.02	2,011,271

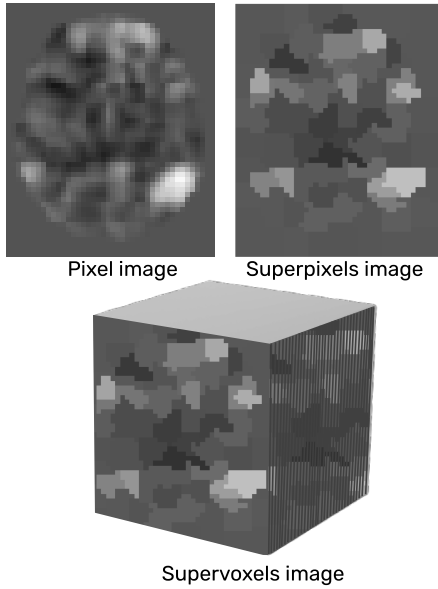


Figure 3: Pixel image segmentation into superpixels and supervoxels.

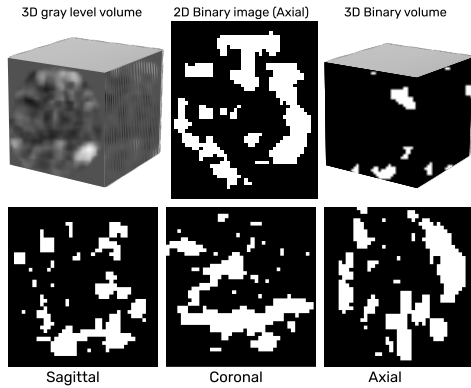


Figure 4: fMRI (LANG network) image spatial transformation.

When using the dimension reduction from 3D to 2D or from grey level to binary images, we observe performance drop as provided in Tab. 1. This suggests that, there is information in the gray level distribution and the 3D shape of the network which are not preserved via the simple spatial dimension reduction tested. By contrast, the values in Tab. 2 represent the functional brain network classification results with CNN model using pixels, superpixels and supervoxels data respectively. Interestingly the loss of performance is very limited when one reduces the gray levels to the average value of the pixels inside a supervoxel or even a superpixel image. Therefore, despite the spatial dimension reduction tested, the reduction of the number of parameters in the models is so far very limited or negligible. To produce this reduction of the model, we proposed to encode the most promising dimension reduction technique (supervoxels) in a compact way as described in the next section.

Table 2: CNN classification of functional brain networks using superpixels/supervoxels images generated in the segmentation stage of graph encoding process with unhealthy subjects as input data.

Data	Train-Test	Accuracy	Parameters
3D gray level	315-70	$0.75 \pm 0.01$	2,356,807
Superpixels image	315-70	$0.69 \pm 0.02$	2,356,807
Supervoxels image	315-70	$0.73 \pm 0.02$	2,356,807

## 4 Graph encoding

To further benefit from the spatial dimension reduction of the previous section, we investigate the possibility to reduce the complexity of the associated neural networks models with limited reduction of performance on the functional cerebral network recognition. To this purpose, we consider to encode our supervoxelized images into graphs. Commonly in graphs, interacting nodes are connected by edges whose weights can be defined by either temporal connections or anatomical junctions, because, graphs are naturally good at relational organization between entities, which makes them great option for representing the 3D capture of voxelwise signals mapped to a specific region of the brain [7]. Therefore, a possibly efficient representation of these fMRI network activations in images can be tested using a graph relation network, which connects nodes of related regions via graph edges.

To obtain a graph representation of our supervoxels images, we connected the segmented neighboring regions through an edge, and denoted the center of each region as a graph node, segment-wise attributes were encoded as node spatial embeddings. This step was repeated until all neighboring nodes were traversed (see Fig. 5). We implemented this approach using the region adjacency graph technique [8], which simply represents each region of the segment as graph nodes and the link between two touching regions as edge using the provided labels of the segmented regions [9]. From the extracted relative spatial coordinates of each superpixel of our image data via the cartesian function, we computed the node position as edge attribute ( $pos[i] - pos[j]$ ) via k-NN graph transformation.

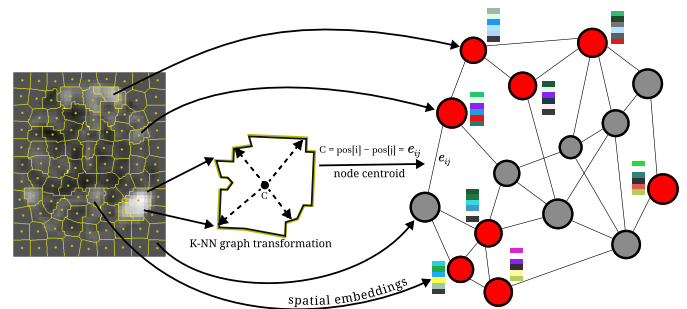


Figure 5: Graph encoding process from superpixels/supervoxels images.

The number of supervoxels was fixed empirically based on the typical size of the activation spots. The resulting graphs

from the encoding stage were observed to be structurally indistinguishable from the connectivity point of view. The contrastive information is expected to stand on the distribution of edge values, which differ from one structural network map to another.

We implemented our method using SplineCNN, a graph neural network which uses a novel type of spline-based convolutional layer for learning [10]. This state-of-the-art GNN is suitable for image-based graph classification task because, it allows the capture of local patterns using spatial relationship between graph nodes by performing global graph pooling. We trained our model parameters with 2 convolutional layers and 2 fully connected output layers with indication of 7 classes in the output layer and a softmax activation. Best results were obtained by training with 2-step learning rate values of  $1e - 3$  for epochs 0 – 200 and  $1e - 5$  for epochs 200 – 500 with early stopping.

For fair comparison with the best result obtained with CNN model in [11], we performed transfer learning during the training of the CNN and GNN models using 80% - 10% - 10% ratio for train-validation-test data split respectively, as well as early stopper with patient set to 10 misses. The performance provided in Tab.3 shows the recorded result from fMRI functional network classification using this transfer learning strategy. Brute transfer indicates the strategy of training directly on healthy data and testing on unhealthy data for both CNN and GNN models. In this cohort, results were compared with values from training and testing on unhealthy data using CNN and GNN model, which provided the 1<sup>st</sup> baseline and 2<sup>nd</sup> baseline values of  $0.75 \pm 0.01$  and  $0.64 \pm 0.03$  respectively, while  $0.78 \pm 0.01$  and  $0.70 \pm 0.01$  were recorded in the transfer learning approach with CNN and GNN respectively. As a consequence, we demonstrate the possibility to obtain a compression of a factor of 26 on the number of model parameters after super-voxelization and graph encoding with only a reduction of 8%.

Table 3: Transfer learning classification with CNN and GNN models.

Data	Train-Test	Accuracy	Parameters
CNN brute-transfer	healthy-unhealthy	$0.75 \pm 0.01$	2,356,807
CNN fMRI healthy (pretrained)	unhealthy-unhealthy	$0.78 \pm 0.01$	2,356,807
GNN brute-transfer	healthy-unhealthy	$0.67 \pm 0.02$	90,855
GNN fMRI healthy (pretrained)	unhealthy-unhealthy	$0.70 \pm 0.01$	90,855

## 5 Conclusion

In this study, we investigated ways to reduce the complexity of end-to-end machine learning models based on convolutional neural networks for the automatic recognition of functional cerebral networks via resting-state fMRI data. A compaction of the activation maps into superpixels or supervoxels shows limited impact on the classification performance. We emphasize the anticipated influence of our 3D multi-channel

images in model parameters, which motivates exploration of a dimension reduction technique before introducing the graph encoding technique. Model evaluation based on spatial dimension reduction was done to investigate its minimal influence in reducing our model parameter. However, this stage was important towards more efficient data encoding (graph structure), which was later shown to have significantly reduced the model parameter. Our initial encoding effort produces a compression of a factor  $26 \times$  where associated reduction in performance was observed at only 8%.

The effort to reduce the complexity of the models was concentrated on the encoding approach of our fMRI data. It would naturally be interesting to couple such effort with investigation on the architecture of the models [12, 13].

## References

- [1] L. Ismaila, P. Rasti, J.-M. Lemée, and D. Rousseau, “Self-supervised learning for functional brain networks identification in fmri from healthy to unhealthy patients,” in *16th International Conference on signal image technology & internet based systems-sitis 2022*. IEEE, 2022.
- [2] L. Ismaila, J.-M. Lemée, D. Rousseau, and P. Rasti, “Deep learning pour la classification automatique de réseaux cérébraux fonctionnels par irmf de repos,” *Journal of Neuroradiology*, vol. 49, no. 2, p. 119, 2022.
- [3] G. Habib and S. Qureshi, “Optimization and acceleration of convolutional neural networks: A survey,” *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 7, pp. 4244–4268, 2022.
- [4] J.-M. Lemée, D. H. Berro, F. Bernard, E. Chinier, L.-M. Leiber, P. Menei, and A. Ter Minassian, “Resting-state functional magnetic resonance imaging versus task-based activity for language mapping and correlation with peri-operative cortical mapping,” *Brain and Behavior*, vol. 9, no. 10, p. e01362, 2019.
- [5] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “Slic superpixels compared to state-of-the-art superpixel methods,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [6] A. Bakkari and A. Fabijańska, “Features determination from super-voxels obtained with relative linear interactive clustering,” in *IPC*, vol. 21, no. 3, 2016, pp. 69–80.
- [7] D. Ahmedt-Aristizabal, M. A. Armin, S. Denman, C. Fookes, and L. Petersson, “Graph-based deep learning for medical diagnosis and analysis: past, present and future,” *Sensors*, vol. 21, no. 14, p. 4758, 2021.
- [8] A. Hagberg, P. Swart, and D. S Chult, “Exploring network structure, dynamics, and function using networkx,” Los

Alamos National Lab.(LANL), Los Alamos, NM (United States), Tech. Rep., 2008.

- [9] L. Wu, P. Cui, J. Pei, L. Zhao, and X. Guo, “Graph neural networks: foundation, frontiers and applications,” in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2022, pp. 4840–4841.
- [10] M. Fey and J. E. Lenssen, “Fast graph representation learning with pytorch geometric,” *arXiv preprint arXiv:1903.02428*, 2019.
- [11] L. E. Ismaila, P. Rasti, F. Bernard, M. Labriffe, P. Menei, A. T. Minassian, D. Rousseau, and J.-M. Lemée, “Transfer learning from healthy to unhealthy patients for the automated classification of functional brain networks in fmri,” *Applied Sciences*, vol. 12, no. 14, p. 6925, 2022.
- [12] R. Sanchez-Iborra and A. F. Skarmeta, “Tinyml-enabled frugal smart objects: Challenges and opportunities,” *IEEE Circuits and Systems Magazine*, vol. 20, no. 3, pp. 4–18, 2020.
- [13] S. Guo and Q. Zhou, *Machine Learning on Commodity Tiny Devices: Theory and Practice*. CRC Press, 2022.