





# RF-Isolation: A Novel Representation of Structural Connectivity Networks for Multiple Sclerosis Classification

Antonella Mensi<sup>1</sup> , Simona Schiavi<sup>1,2</sup> , Maria Petracca<sup>3,4</sup>,  
Nicole Graziano<sup>4</sup>, Alessandro Daducci<sup>1</sup>, Matilde Inglese<sup>2,4</sup>,  
and Manuele Bicego<sup>1</sup>

<sup>1</sup> Department of Computer Science, University of Verona, Verona, Italy  
`antonella.mensi@univr.it`

<sup>2</sup> Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), University of Genoa, Genoa, Italy

<sup>3</sup> Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy

<sup>4</sup> Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

**Abstract.** Magnetic Resonance Imaging (MRI) is one of the tools used to identify structural and functional changes caused by multiple sclerosis, and by processing MR images, connectivity networks can be obtained. The analysis of structural connectivity networks of multiple sclerosis patients usually employs network-derived metrics, which are computed independently for each subject. We propose a novel representation of connectivity networks that is extracted from a model trained on the whole multiple sclerosis population: RF-Isolation. RF-Isolation is a vector encoding the disconnection of each region of interest with respect to all other regions. This feature can be easily captured by isolation-based outlier detection methods. We therefore reformulate the task as an outlier detection problem and propose a novel approach, called MS-ProxIF, based on a variant of Isolation Forest, a Random Forest-based outlier detection system, from which the representation is extracted. We test the representation via a set of classification experiments, involving 79 subjects, 55 of which suffer from multiple sclerosis. In particular, we compare favourably to the most used network-derived metrics in multiple sclerosis.

**Keywords:** Multiple sclerosis · Structural connectivity network · Microstructure informed tractography · Proximity isolation forest

## 1 Introduction

Multiple Sclerosis (MS) is a chronic autoimmune disease of the central nervous system causing demyelination and neurodegeneration [1]. It is usually diagnosed and followed up via the analysis of Magnetic Resonance Imaging (MRI), which is sensitive to demyelination, i.e. lesions affecting white matter tracts and causing disconnection of grey matter regions.

The aim of this paper is to discern MS patients from healthy subjects, starting from their quantitative connectivity networks estimated with MRI. A connectivity network encodes the connectivity strength between each pair of brain regions of interest (ROIs). An interesting technique, used to compare MS subjects to a healthy cohort, consists of analyzing the distribution of network-derived metrics. These measures capture the importance in terms of connectivity of either the whole network or a subnetwork composed by fewer ROIs. The rationale behind these methodologies is that in connectivity networks, because of the presence of lesions, the connectivity strength between the ROIs involved in a lesion is expected to be lower in an MS subject than in a healthy one. Although the relevance of network-derived measures in the study of MS has been already investigated [11, 14, 15], there are some issues that would benefit from methodological improvement. In particular, all these measures are subject-wise: they are computed independently for each subject using only the connectivity network of the subject under analysis. In other words, they tend to capture how MS behaves in a specific subject, rather than the global nature of the disease.

In this paper, we propose a novel representation of MS connectivity networks, called *RF-Isolation*. The main characteristics of *RF-Isolation* are: i) it measures the disconnection of each ROI with respect to all other ROIs; ii) the problem is reformulated as an outlier detection task, hence the disconnection of a ROI is represented via an outlierness score, extracted by using a Random Forest (RF)-based outlier detector model; iii) it is disease-wise, i.e. it is extracted for all subjects, MS and healthy, from a model built using the entire MS population.

We assessed the suitability and robustness of the proposed approach via several classification experiments, performed on a cohort of 79 subjects, 55 of which suffer from MS. We also compared to standard network-derived metrics.

## 2 Materials and Methods

### 2.1 Study Population

The dataset we employ consists of 79 subjects<sup>1</sup>: 55 suffering from MS and 24 healthy controls. MS subjects had to meet several inclusion criteria: for example, among others [15], their Expanded Disability Status Scale score had to be  $\leq 7$  and they had to satisfy the diagnostic McDonald criteria. There were also several exclusion criteria, such as: the presence of any major systemic condition, pregnancy, and addiction to drugs/alcohol, among others [15]. Further, they underwent a clinical examination within a week from the MRI scan. All subjects have signed a written informed consent prior to the beginning of the whole study, as the Declaration of Helsinki states. The Institutional Review Board of the Icahn School of Medicine at Mount Sinai approved the protocol.

---

<sup>1</sup> The dataset has been collected at the Mount Sinai Hospital of New York (US) by the group of Matilde Inglese. The dataset is not publicly available.

## 2.2 MRI Acquisition and Processing

The MRI acquisition protocol was the same for all subjects: a Siemens Skyra 3T scanner (Siemens, Erlangen, Germany) with a 32-channels head coil was used –further technical details can be found in [15]. As to the processing of the brain images, the first part of the pipeline illustrated in Fig. 1 of [14] was followed. Briefly, images were segmented obtaining a cortical parcellation in 85 grey matter ROIs using the Desikan-Killiany atlas [6]. At the same time, the tractography was computed using a probabilistic algorithm. The next crucial step was the application of the COMMIT (Convex Optimization Modeling for Microstructure Informed Tractography) framework [4] to obtain quantitative structural connectivity networks that better reflect the white matter tissue microstructure.

The final connectivity networks, one for each subject, were obtained by combining the COMMIT-weighted tractogram with the segmented grey matter. Each connectivity network is composed by 85 brain regions of interest of grey matter, encoded as nodes; the connectivity strength between each pair of ROIs is encoded as an undirected edge between the involved ROIs. Please note that in the rest of the paper, we will use the terms ROIs and brain regions interchangeably.

## 2.3 RF-Isolation Extraction

The proposed methodology, *RF-Isolation*, is based on three assumptions:

- i *RF-Isolation* is a vector of length equal to the number of established brain regions. Each feature of the vector encodes the degree of disconnection of one ROI with respect to all other ROIs. Indeed, disconnections in the brain are a cornerstone of MS.
- ii To encode such characteristic, we map the problem to the outlier detection context. Outliers are objects which do not conform to the rest of the data: they are few and different from the former. We can easily interpret a ROI with a high level of disconnection as an outlier. Therefore, the disconnection level of a ROI can be measured by quantifying its “outlierness”.
- iii The last assumption is that to extract a meaningful representation for MS, we should build the model using only subjects suffering from MS. Using only the healthy cohort, or both populations, would lead to a model in which the identification of disconnected ROIs is presumably more difficult.

In detail, our approach is based on two steps:

1. Train a model using the *entire MS population*.
2. Given a subject, we can obtain its *RF-Isolation* vector by employing the model from Step 1.

Before thoroughly describing each step, a note must be made on how the adopted model was chosen: our reasoning starts from Isolation Forest (iForest) [8], an RF-based methodology that is also one of the most successful outlier detectors. The aim of iForest is to separate each object from the rest of the data, relying on the

principle of isolation. Outliers, due to their nature, are likely to get isolated early in a tree, i.e. they have a high isolation capability. The latter is used to quantify the anomaly score, i.e. outlieriness degree, of an object. Unfortunately, iForests work only with vectorial data, whereas our starting point is a connectivity network, which can be seen as a similarity matrix. In detail, the connection strength between two ROIs can be interpreted as a similarity value, i.e. the stronger the connection, the higher the similarity. Thus, we based our approach on Proximity Isolation Forest (ProxIF) [9], an RF and isolation-based outlier detection methodology that works with all types of data for which a proximity measure can be defined. In detail, we extend ProxIF and adapt it for this applicative context: we denote the obtained model as *MS-ProxIF*.

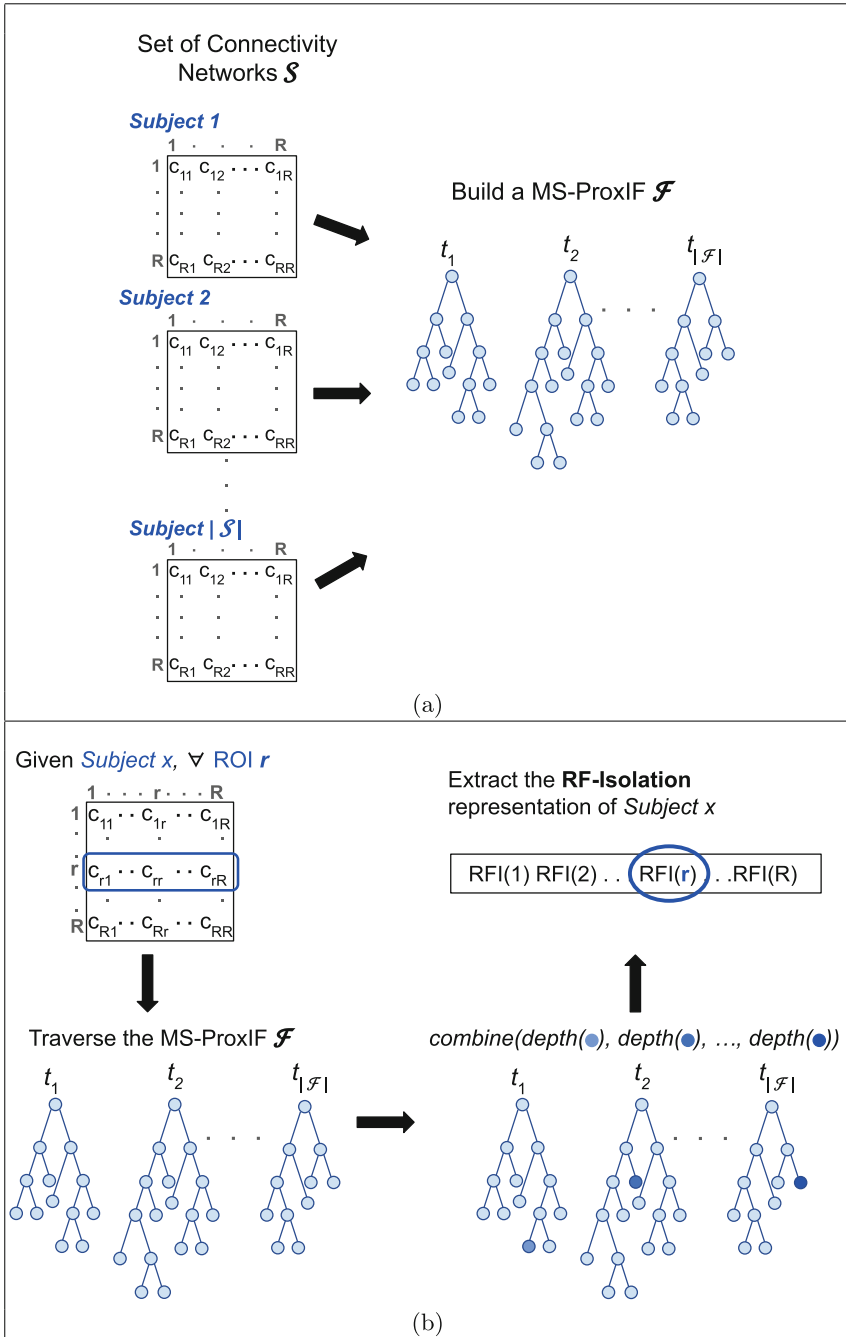
Figure 1 depicts the pipeline of the proposed approach. In Fig. 1(a) we illustrate the building procedure of an MS-ProxIF model, whereas in Fig. 1(b) we use the MS-ProxIF model from Fig. 1(a) to extract the *RF-Isolation* representation. Each step is thoroughly described in the next two subsections.

**Step 1.** The first step, i.e. the building procedure of MS-ProxIF, is the most complex. A ProxIF is an ensemble learner composed by several randomized decision trees, called Proximity Isolation Trees (ProxITs). Each ProxIT is built using a similarity matrix encoding pairwise similarities, which in our context, corresponds to building an MS-ProxIT using one connectivity network. Actually, differently from ProxIT, each MS-ProxIT can be built using several connectivity networks, each representing a different subject. This extension is crucial, since each MS-ProxIT, rather than capturing subject-wise characteristics, can retrieve disease-specific information. Therefore, an MS-ProxIT is built using a random subset of both ROIs and connectivity networks, each drawn without replacement.

Before describing how to build an MS-ProxIT, we must define how a ROI  $x$  traverses it. The traversal procedure is recursive and describes whether a ROI in a node  $n$  should go to the left child  $n_L$  or to the right one  $n_R$ . In detail, in [9] two traversal modalities are defined:

- Given a node  $n$  we have one prototype  $P$ , i.e. a ROI, and a threshold on the connectivity strength  $\theta \in [\min_{x \in n} \text{connectivity}(x, P), \max_{x \in n} \text{connectivity}(x, P)]$ , if  $\text{connectivity}(x, P) > \theta$  then  $x \rightarrow n_L$ , otherwise  $x \rightarrow n_R$ . In other words,  $x$  ends up in the left child, along with the prototype, only if their connection is strong enough.
- In a node  $n$  two different ROIs,  $P_L$  and  $P_R$  are chosen as prototypes, which respectively represent the putative left and right child. If  $\text{connectivity}(x, P_L) > \text{connectivity}(x, P_R)$  then  $x \rightarrow n_L$ , otherwise  $x \rightarrow n_R$ . In other words, the ROI  $x$  ends up in  $n_L$  if its connection to  $P_L$  is stronger than its connection to  $P_R$ .

In our context, the first traversal modality seems more suitable. In detail, when using only one prototype, by analyzing how strongly connected is the ROI  $x$  to the ROI  $P$ , we obtain a partial contribution to the total disconnection degree of  $x$ , which computation is the final aim of the proposed methodology.



**Fig. 1.** Pipeline of the proposed approach: (a) building an MS-ProxIF; (b) extracting the RF-Isolation vector for a subject  $x$ . *Combine* stands for the function computing the anomaly score at forest level via aggregation of the tree scores.

Instead, from the second traversal modality, we can only infer to which prototype the ROI  $x$  has a stronger connection, which is not what we are seeking.

The aim behind the building procedure of a ProxIT, is to obtain a tree structure where splitting  $n$  generates two child nodes such that: each child contains objects highly similar to each other but dissimilar to the objects in the sibling node. In our context, the aim is to find one child node that contains many ROIs strongly connected with each other, and the other child that contains few ROIs. The former will be the root of a big subtree, and thus the contained ROIs will be isolated after a long procedure. The converse holds for the ROIs in the smaller sibling: they are likely to be isolated soon, i.e. of being disconnected. Ideally we would like a split of a node  $n$  to generate a node containing only inliers, highly connected ROIs, and one containing only outliers, disconnected ROIs.

To build a tree, we have to define a learning strategy, i.e. how to define on a node  $n$  the test which induces its splitting into two child nodes. In an MS-ProxIT, it consists of choosing: one prototype ROI  $P$ , a threshold  $\theta$  on the connectivity strength with respect to  $P$  (or two prototypes  $P_L$  and  $P_R$ <sup>2</sup>), and in some cases, additionally with respect to [9], a connectivity network. Indeed, an MS-ProxIT can be built with one or several connectivity networks. In the first case, the learning algorithm is identical to [9]: given  $n$ ,  $P$  and  $\theta$  we split the ROIs in  $n$ , based on how strongly connected they are to  $P$ . In the latter scenario, instead, since the MS-ProxIT is built using multiple connectivity networks, the test on node  $n$  is characterized by  $P$ ,  $\theta$  and one connectivity network  $cn$ . In other words, to split the ROIs in  $n$ , we evaluate how strongly connected they are to  $P$  in the chosen connectivity network  $cn$ . It is important to note that: i) only ROIs traverse the nodes; ii) the connectivity network  $cn$  is needed *only during* the tree building procedure. Therefore, when the latter has ended, we can discard the information about which  $cn$  was used to partition the ROIs in  $n$ .

The other crucial step of the building procedure consists of choosing the best test for each node  $n$ : the choice can be either completely random or based on an optimization procedure. This step is independent of the number of connectivity networks used to build the tree.

In detail, to the already extensive pool of training criteria proposed in [9], i.e. how to choose the best test, we add two additional ones, *O-1PRD* and *O-2PRD*. These criteria are an adaptation to the outlier detection context of the *RényiD* [2] criterion proposed for clustering, i.e. the best split is the one maximizing the divergence in terms of information (quantified by the Rényi entropy) between the two child nodes. This concept is relevant also for outlier detection, since finding the two child nodes conveying the highest amount of different information, corresponds to ideally separating inliers from outliers. In detail, to define these criteria, we estimate the Rényi divergence between the child nodes using an estimator [10] which employs only information related to the K-Nearest Neighbors (KNNs) of the objects. Formally, given a set  $A$  of  $N$

---

<sup>2</sup> For the sake of clarity the remainder of the explanation will refer to tests characterized by  $P$  and  $\theta$ , but an analogous reasoning would hold if the test consisted of choosing two prototypes  $P_L$  and  $P_R$ .

objects and a set  $B$  of  $M$  objects, the estimation of the Rényi divergence of order  $\alpha$  of  $A$  from  $B$  is computed as follows:

$$RD(A, B) = \frac{1}{\alpha - 1} \log \left[ \frac{\frac{M}{N}^\alpha}{M} \sum_{i=1}^M \left( \frac{N_i}{M_i + 1} \right)^\alpha \right] \quad (1)$$

where  $N_i$  and  $M_i$  are the number of KNNs of  $b_i \in B$  that respectively belong to  $A$  and  $B$ . In other words,  $A$  will diverge more from  $B$ , if few of its objects are KNNs of objects in  $B$ . In our context,  $A$  and  $B$  are the child nodes  $n_L$  and  $n_R$ , and to measure their divergence, we have to compute both  $RD(n_L, n_R)$  and  $RD(n_R, n_L)$  –and average their results– since the divergence is not symmetric. Please note that *O-1PRD* and *O-2PRD* differ only because the former evaluates tests defined by  $P$  and  $\theta$ , whereas *O-2PRD* evaluates tests defined by two prototypes  $P_L$  and  $P_R$ . For additional information related to the learning phase, see [9].

**Step 2.** Given a built MS-ProxIT, we can extract, for a given subject and a given ROI  $x$ , its anomaly score. Due to their outlier-like nature, disconnected ROIs are more likely to be isolated sooner in the tree, and therefore we would like to assign them a higher anomaly score. To assign such score, we make the ROI  $x$  traverse the tree, following one of the two modalities previously presented, depending on the training criterion used to build  $t$  (one or two prototypes). It is important to highlight that traversing  $t$  is independent of which and how many connectivity networks were used to build the MS-ProxIT. In other words, the evaluation of whether a ROI  $x$  should follow the left or right edge depends exclusively on the connectivity network of the subject that is traversing  $t$ .

After traversing  $t$ , we can recover the anomaly score of the ROI  $x$ , which is a function of the depth of the reached leaf. In detail, the smaller the depth, the higher the anomaly score, i.e. the more likely the ROI  $x$  is highly disconnected. To obtain the anomaly score of  $x$  at forest level, we have to make  $x$  traverse each MS-ProxIT composing the MS-ProxIF: the final score is a function of the average depth of the reached leaves –see [8] for details on the formula.

By repeating this procedure for all ROIs of a subject, we obtain the *RF-Isolation* vector, a novel representation where each feature represents the degree of disconnection of a ROI as computed from MS-ProxIF.

Please note that an *RF-Isolation* vector, extracted from an MS-ProxIF – which we recall is trained on the MS population– is an adequate representation also for the healthy cohort. Indeed, given a ROI disconnected only in MS subjects, in a healthy patient the same ROI is likely to traverse a longer path, i.e. get a lower anomaly score, since its connectivity to other ROIs is probably higher.

## 2.4 Classification Analysis

To evaluate the suitability and robustness of the proposed representation, we made two different analyses: the first studies the MS-ProxIF model, whereas the second one compares *RF-Isolation* to standard network-derived metrics.

Both analyses are based on the classification of the 79 subjects as either MS or healthy, following a Leave One Out protocol (LOO): each classifier is trained 79 times on a dataset of 78 subjects and tested on the left out subject. The classifiers we employed are: linear Support Vector Machines (SVM), K-Nearest Neighbor (KNN) and Random Forest (RF). We measured the classification performances using the Matthews Correlation Coefficient [3] (MCC) which takes into account all four entries of the confusion matrix and thus returns a high value only if both classes are well identified. In detail, we employ the normalized MCC, which ranges in  $[0, 1]$ .

To increase the robustness of the proposed methodology, we repeat 5 times the whole procedure, i.e. from building the MS-ProxIF to the classification step.

**Analysis of the MS-ProxIF Model.** The first analysis aims at finding the most suitable variant of the MS-ProxIF model for the task of extracting RF-Isolation—and subsequently of discriminating MS subjects from the healthy cohort. In detail, we generated many different MS-ProxIF models, from which we extracted different *RF-Isolation* vectors, by varying the following parameters:

- Number of trees in a forest:  $T \in \{50, 100, 200\}$ .
- Number of ROIs used to build each tree:  $S \in \{50\%, 75\%, 100\%\}$ .
- Number of connectivity networks used to build each tree:  $C \in \{1, 5, 10, 20, All\}$  where *All* consists of using all MS subjects to build each tree.
- The training criterion. In detail, we study 6 criteria: *R-1P* and *R-2P* which split data randomly, chosen because of the pervasive success of random variants in [8]; *O-1PS<sub>D</sub>* and *O-2PS<sub>D</sub>*, the best variant according to [9]; and *O-1PRD* and *O-2PRD*, the novel variants proposed in Sect. 2.3.
- The maximum depth that each tree can reach:  $D \in \{\log_2(S), S - 1\}$ .

We analyze the behaviour of each of the above parameters via a statistical analysis of the classification results. The input of each analysis is the set of MCC values averaged across the 5 iterations of the proposed approach and all parameters except: the 6 training criteria, the 3 classifiers (SVM, KNN, RF) and the parameter under analysis. For example, if we were to analyze  $T$ , the input would be made of 54 MCC values since: we average across the iterations and all values of  $D$ ,  $S$  and  $C$ ; we have 3 values of  $T$ , 3 classifiers and 6 training criteria.

As to the adopted statistical procedure, when analyzing  $D$ , since it can assume only two values, we perform a Wilcoxon signed-rank test, whereas for all other parameters, we carry out a Friedman test followed by a post-hoc Nemenyi test. Indeed, the Friedman test is adopted if we have to assess whether there is a global significant difference among more than two methods, in our scenario represented by the different values a parameter can take. Following the Friedman test, we perform a Nemenyi test that employs a critical value to find out which pairs of methods are statistically different. We use a critical difference (CD) diagram [5] to visualize the results of these tests. A CD diagram consists of a line where methods are represented from left to right based on their rank, from worst to best. If two methods are not significantly different, a red line connects them. The significance level is set to  $\alpha = 0.05$  for all tests.



**Comparison to Standard Network Measures.** The second analysis aims at understanding whether *RF-Isolation* has a higher capability of correctly identifying MS subjects compared to standard network-derived measures. In detail, from each subject’s connectivity network, we extracted the following metrics using the Brain Connectivity Toolbox [13]<sup>3</sup>:

- *Node Strength*: It is the sum of the weights of all edges connected to the node.
- *Local Efficiency*: It is the average inverse shortest path length in the neighborhood of the node; a stronger connection leads to a stronger contribution.
- *Assortativity*: It is a correlation coefficient based on the Node Strength. A positive value indicates that nodes tend to link together with nodes which Node Strength is similar.
- *Clustering Coefficient*: It is the sum of the clustering coefficients of all nodes. The Clustering Coefficient of a node consists of the proportion of its neighbors which are neighbors of each other.
- *Density*: It is the ratio between the number of existing connections to the number of possible connections.
- *Global Efficiency*: It is the average inverse shortest path length in the network.
- *Mean Strength*: It is the average Node Strength computed across all nodes.
- *Modularity*: It is a statistic describing to which degree the network can be partitioned into disjoint sets of nodes, such that within each set the number of edges maximized, and the number of edges between different groups is minimized.

Node Strength and Local Efficiency are local measures, i.e. analogously to *RF-Isolation*, they have one feature for each ROI. Instead, all the remaining metrics are global measures, i.e. one value describes a connectivity network. These measures were chosen because of their significance with respect to MS, as shown in previous works ([15] and references therein). As mentioned in Sect. 1, all the listed metrics are subject-wise, being computed independently for each subject.

### 3 Results

#### 3.1 Analysis of the MS-ProxIF Model

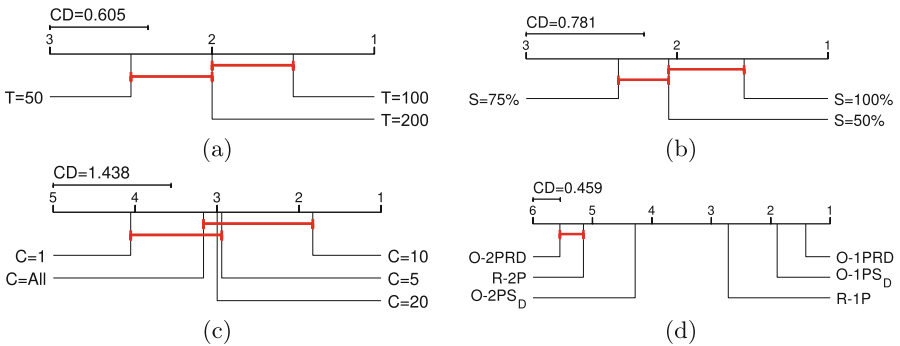
Table 1 depicts the results of the analysis of the depth  $D$ : we report the mean rank of both values of  $D$ , and the p-value output by the test, in **bold** if the null hypothesis is rejected. We can assess that using a smaller depth,  $D = \log_2(S)$ , is the best significant option. The latter is a common choice in the context of isolation-based approaches [8], and it also represents the average tree height [7]. The second parameter we analyze is the forest size  $T$ , which results are depicted via the CD diagram in Fig. 2(a): even though  $T = 100$  is first-ranked, it is comparable to  $T = 200$ . Nevertheless,  $T = 100$  remains a wiser option

<sup>3</sup> For more thorough descriptions, please refer to <https://sites.google.com/site/bctnet/list-of-measures>.

**Table 1.** Results of the Wilcoxon signed-rank test comparing the two options for the maximum depth  $D$ .

Mean rank		p-value
$D = \log_2(S)$	$D = S-1$	
1.222	1.778	<b>0.0074</b>

from a computational perspective, and it is also a common choice for RF-based methodologies [8, 12]. Then we analyze the number of ROIs  $S$  used to build each tree, which is expressed in percentage 50% = 43 ROIs, 75% = 64 ROIs, and 100% = 85 ROIs. From the CD diagram depicted in Fig. 2(b) we can observe that using all ROIs is the best choice, even though it is comparable to  $S = 50\%$ . One of the most interesting parameters is  $C$ , the number of connectivity networks used to build each tree. The results of the statistical analysis, shown in Fig. 2(c), assess that  $C = 10$  is the best option, even though it is comparable to all other values of  $C$ , except for  $C = 1$ , the last ranked. Results confirm the usefulness of employing multiple connectivity networks, i.e.  $C > 1$ , to build each tree, since it leads to a more informative *RF-Isolation* representation. The last analysis compares the different training criteria: the CD diagram in Fig. 2(d) assesses that *O-1PRD* is the best choice. Further, training criteria based on two prototypes are all ranked significantly worse than all criteria using one prototype, making the latter a better choice, as hypothesized.

**Fig. 2.** CD diagram comparing the different options for: (a) the forest size  $T$ ; (b) the number of ROIs  $S$  used to build each tree; (c) the number of connectivity networks  $C$  used to build each tree; (d) the adopted training criterion.

Summarizing, we can conclude that even though we can extract a robust *RF-Isolation* vector from several parametrizations, we can identify one that is the most suitable. In detail, we should set the parameters of *MS-ProxIF* as follows:  $D = \log_2(S)$ ,  $S = 100\%$ ,  $C = 10$ ,  $T = 100$ , and the training scheme to *O-1PRD*.

### 3.2 Comparison to Standard Network Measures

In Table 2 we report the comparison, in terms of MCC, between standard network-derived measures, and the proposed approach. As to *RF-Isolation*, the MCC has been averaged across the iterations, and the parameters of the underlying MS-ProxIF model have been set according to the previous analysis. We can observe that *RF-Isolation* shows better performances than all the other measures independently of the employed classifier. The only exception is when we compare to Local Efficiency using KNN: however, the difference is rather small with respect to the improvements brought by *RF-Isolation*. To validate the improvement, we also report the standard errors of the mean.

Obtained results suggest that *RF-Isolation*, a disease-wise representation extracted from a trained model, represents a valid alternative to standard metrics for connectivity networks, which are instead subject-wise.

**Table 2.** Comparison of RF-Isolation with standard network-derived metrics.

Representation	MCC		
	SVM	KNN	RF
<i>RF-Isolation</i>	<b>0.8686 (0.0128)</b>	0.7822 (0.0192)	<b>0.8554 (0.0139)</b>
Local efficiency	0.7997 (0.0180)	<b>0.7843 (0.0190)</b>	0.7799 (0.0193)
Nodal strength	0.8167 (0.0168)	0.6559 (0.0254)	0.8127 (0.0171)
Assortativity	0.5000 (0.0281)	0.5000 (0.0281)	0.4790 (0.0281)
Clustering coefficient	0.5000 (0.0281)	0.5950 (0.0271)	0.6110 (0.0267)
Density	0.5000 (0.0281)	0.4369 (0.0277)	0.6325 (0.0262)
Global efficiency	0.6226 (0.0264)	0.7373 (0.0218)	0.5818 (0.0274)
Mean strength	0.7261 (0.0224)	0.6760 (0.0246)	0.6102 (0.0268)
Modularity	0.5812 (0.0274)	0.7487 (0.0212)	0.6392 (0.0259)

## 4 Conclusion

We proposed a novel metric of quantitative connectomes to analyse MS patients via classification. Differently from other graph-based metrics, the novel representation, *RF-Isolation*, is more descriptive from a disease point of view. Indeed, to extract it, we trained a model on the entire MS population, from which we derived the *RF-Isolation* vector for all subjects. The suitability of the approach is confirmed by the experimental analyses. Future work includes analyzing the link between *RF-Isolation*, and the most involved ROIs in the MS lesion process. In addition, when a bigger dataset will be available, we will test *RF-Isolation* for classifying different subtypes of MS, which is nowadays clinically challenging. Lastly, we could also adapt the proposed methodology to study other neurodegenerative diseases, e.g. Amyotrophic Lateral Sclerosis.

**Acknowledgments.** The authors declare that there is no conflict of interest. This work was partly supported by grants from NMSS (RG 5120A3/1) and Teva Neuroscience (CNS-2014-221).

## References

1. Bester, M., Petracca, M., Inglese, M.: Neuroimaging of multiple sclerosis, acute disseminated encephalomyelitis, and other demyelinating diseases. In: *Seminars in Roentgenology*, vol. 49, pp. 76–85 (2013). <https://doi.org/10.1053/j.ro.2013.09.002>
2. Bicego, M.: Dissimilarity random forest clustering. In: 2020 IEEE ICDM, pp. 936–941 (2020). <https://doi.org/10.1109/ICDM50108.2020.00105>
3. Chicco, D., Jurman, G.: The advantages of the Matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. *BMC Genomics* **21**(1), 1–13 (2020). <https://doi.org/10.1186/s12864-019-6413-7>
4. Daducci, A., Dal Palù, A., Lemkaddem, A., Thiran, J.P.: COMMIT: convex optimization modeling for microstructure informed tractography. *IEEE Trans. Med. Imaging* **34**(1), 246–257 (2014). <https://doi.org/10.1109/TMI.2014.2352414>
5. Demšar, J.: Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **7**, 1–30 (2006)
6. Desikan, R.S., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**(3), 968–980 (2006). <https://doi.org/10.1016/j.neuroimage.2006.01.021>
7. Knuth, D.E.: The art of computer programming. *Sorting Search*. 3, Ch–6 (1973)
8. Liu, F.T., Ting, K.M., Zhou, Z.H.: Isolation-based anomaly detection. *ACM Trans. Knowl. Discov. Data* **6**(1), 1–39 (2012). <https://doi.org/10.1145/2133360.2133363>
9. Mensi, A., Bicego, M., Tax, D.M.: Proximity isolation forests. In: 2020 25th ICPR, pp. 8021–8028. IEEE (2021). <https://doi.org/10.1109/ICPR48806.2021.9412322>
10. Noshad, M., Moon, K.R., Sekeh, S.Y., Hero, A.O.: Direct estimation of information divergence using nearest neighbor ratios. In: 2017 IEEE International Symposium on Information Theory (ISIT), pp. 903–907. IEEE (2017)
11. Pagani, E., et al.: Structural connectivity in multiple sclerosis and modeling of disconnection. *Mult. Scler. J.* **26**(2), 220–232 (2020). <https://doi.org/10.1177/1352458518820759>
12. Probst, P., Boulesteix, A.L.: To tune or not to tune the number of trees in random forest. *J. Mach. Learn. Res.* **18**(1), 6673–6690 (2017)
13. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**(3), 1059–1069 (2010). <https://doi.org/10.1016/j.neuroimage.2009.10.003>
14. Schiavi, S., et al.: Classification of multiple sclerosis patients based on structural disconnection: a robust feature selection approach. *J. Neuroimaging* (2022). <https://doi.org/10.1111/jon.12991>
15. Schiavi, S., et al.: Sensory-motor network topology in multiple sclerosis: structural connectivity analysis accounting for intrinsic density discrepancy. *Hum. Brain Mapp.* **41**(11), 2951–2963 (2020). <https://doi.org/10.1002/hbm.24989>