

26th International Conference on Knowledge-Based and  
Intelligent Information & Engineering Systems (KES 2022)

# Machine learning techniques for MRI feature-based detection of frontotemporal lobar degeneration

Tatiana Pilipenko\*, Alessandro Gnutti, Andrea Silvestri, Ivan Serina, Riccardo Leonardi

*University of Brescia, Department of Information Engineering, via Branze 38, Brescia 25123, Italy*

---

## Abstract

Making a diagnosis of neurodegenerative diseases at an early stage is one of the most significant challenges of modern neuroscience. Although this family of diseases remains without a cure, the effectiveness of their medical treatment largely relies on the timing of their detection. For certain groups of diseases, such as Fronto-Temporal Dementia (FTD), trained professionals can effectively reach a correct diagnosis through the visual analysis of Magnetic Resonance Imaging, in its functional (fMRI) or raw (MRI) version. However, this operation is time-consuming and may be subject to personal interpretation. In this paper, we explore the performance of a group of machine learning algorithms to formulate a correct FTD diagnosis, in order to provide medical professionals with a supporting tool. The dataset consists of MRI data acquired on 30 subjects, and the experiments are carried out by investigating different fMRI techniques based on a Multi-Voxel Pattern Analysis (MVPA) approach. The results obtained show high accuracy in identifying FTD in elderly patients when Support Vector Machine and Random Forest techniques are used, with outcomes varying based on the fMRI methods.

© 2022 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

**Keywords:** Frontotemporal Dementia; Multi-Voxel Pattern Analysis; Machine Learning; Support Vector Machines; Random Forest.

---

## 1. Introduction

Frontotemporal Dementia (FTD) is a neurodegenerative disorder that is believed to hit 60,000 people only in the United States per year [12]. FTD has distinct clinical phenotypes associated with multiple neuropathologic entities, and it encompasses clinical disorders that include changes in behaviour, language, executive control and often motor symptoms [15]. FTD has no cure and, as of today, there exist only supportive treatments that are able to slow down its course. Therefore, it is crucial to detect the presence of this neuropathology at its early stage of development in order to improve the quality of life of patients and their families.

---

\* Corresponding author.

*E-mail address:* [t.pilipenko@unibs.it](mailto:t.pilipenko@unibs.it)

Magnetic resonance imaging (MRI) is a well-established method for FTD diagnosis that can give a full picture of the neurodegenerative condition in a patient. The important role of MRI is not only recognised in its use to diagnose, but also in its application to exclude other pathologies. Furthermore, the progress made in the MR methodology allowed the possibility of FTD detection in its early stages. In general, a medical expert observes MRI data and evaluates the patient's condition. However, this operation is time-consuming and may be subject to personal interpretation. As a matter of fact, the study of methods designed to automatically elaborate the information contained in MR images is nowadays a very active research field.

In this paper, we investigate the performance of some artificial intelligence (AI) algorithms for automatic classification of brain images associated to FTD carriers and to healthy subjects, of which Support Vector Machine (SVM) and Random Forest (RF) techniques have returned the best results. AI is a field that has been developing in the medical area very rapidly over the recent years. We can find different researches in anaesthesiology [7], nephrology [14], drug design [8] and cardiology [9], but nowadays the most popular field for the implementation of AI in the health system is medical imaging [5, 13, 19, 23]. In this field, the automatic classification of MRI data has become increasingly popular in diagnosing neurodegenerative diseases [4, 11].

Here, we extend the work proposed in [18], in which an SVM model was proposed as classifier and different MRI functional parameters were considered as features. The diagnostic process developed by [18] has been enhanced by performing an accurate hyper-parameter optimization leading to important improvements on the overall reliability of the procedure. Our best result achieves an average accuracy of 80% at discerning FTD patients from healthy subjects.

The benefits of this research can be analyzed from a dual perspective: (i) from a strictly medical point of view, this work may represent an important step towards a future where doctors will be supported by computers in their diagnostic work, while, (ii) from the understanding in FTD, the identification of small patterns in the MR images may open novel research directions, as it reveals micro-patterns that would have been difficult to consider otherwise.

The rest of the paper is organized as follows. In Sec. 2, we list the methods used in our work. In particular, the dataset and the considered MRI functional parameters are presented in Sec. 2.1, while the investigated machine learning (ML) algorithms are described in Sec. 2.2. In addition, the pre-processing stages applied on the dataset and the procedure of the experiments are illustrated in Sec. 2.3 and 2.4, respectively. Results are then presented and discussed in Sec. 3. Specifically, the comparison with Premi's results is shown in Sec. 3.1, while additional experiments are detailed in Sec. 3.2. Conclusions are finally drawn in Sec. 4

## 2. Methodology

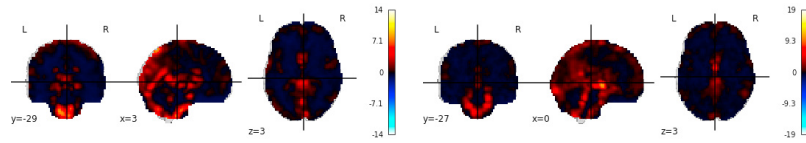
### 2.1. Dataset and features

The dataset was acquired at the Centre for Ageing Brain and Neurodegenerative Disorders, University of Brescia (Brescia, Italy), and consists of MRI data on 30 subjects, divided in 2 classes: 14 patients with FTD carrying GRN Thr272fs mutation (FTD-GNR+) <sup>1</sup> and 16 elderly gender-matched healthy subjects (oHC). Informed consent was obtained from all the subjects or their responsible guardians involved in the research. The work was organized in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Brescia Hospital.

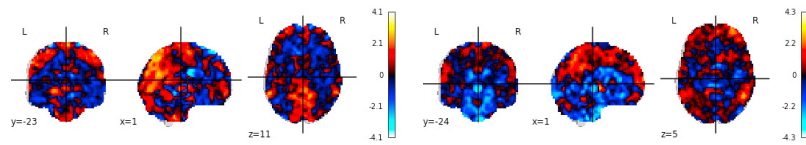
Different functional parameters have been applied to the MRI dataset, such as Amplitude of Low Frequency Fluctuations (ALFF), Fractional Amplitude of Low Frequency Fluctuations (fALFF), Degree Centrality (DC), Regional Homogeneity (ReHo) and Voxel-Mirrored Homotopic Connectivity (VMHC). Fig. 1 shows an example of fMRI image for each functional parameter and for each class.

The ALFF and fALFF measures are related to each other: ALFF is a measure that highlights, within specific regions, the physiological states of a resting brain and its spontaneous neural activity, whereas fALFF is a measure used in order to reduce the sensitivity of ALFF to physiological noise of a brain (e.g., hunger, fatigue, pain, etc.) [25]. VMHC provides an index of functional symmetry of a resting-state brain activity [26], while DC is a measure that ex-

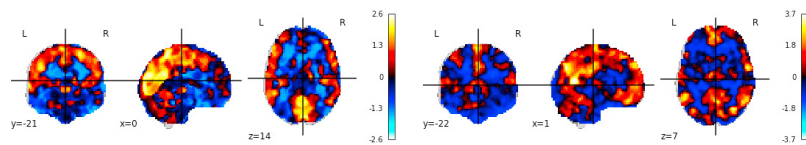
<sup>1</sup> Progranulin is a glycoprotein, encoded in humans by the GRN gene, located on the 17q21 chromosome. A mutation of this gene, named Thr272fs, is responsible for frontotemporal lobar degeneration cases [2]



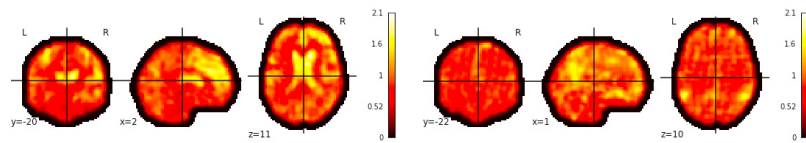
(a) Amplitude of Low Frequency Fluctuations (ALFF).



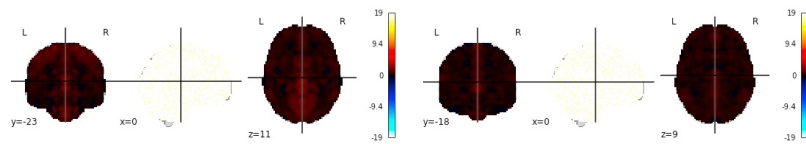
(b) Fractional Amplitude of Low Frequency Fluctuations (fALFF).



(c) Degree Centrality (DC).



(d) Regional Homogeneity (ReHo).



(e) Voxel-Mirrored Homotopic Connectivity (VMHC).

Fig. 1: An example of fMRI image for each functional parameter. On the left the class FTD-GNR+, on the right the class oHC.

tracts information from the single nodes composing the whole-brain network [3]. Finally, ReHo is the time consistency of blood oxygenation in local brain tissue [24].

The dataset used in this work can be requested to the Centre for Ageing Brain and Neurodegenerative Disorders (University of Brescia, Italy) for research purposes.

## 2.2. Proposed classification methods

In this work, different ML techniques have been tested for the classification task. Since the experiments have shown that SVM and RF approaches clearly outperform other algorithms, this paper focuses on them. However, for the sake

of completeness, we remark that also Extra Trees, K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA), Gradient Boosting, and Multilayer Perceptron (MLP) have been tested. The reader can find the complete setup of the experiments in [17]

SVM is a supervised algorithm that can be used for classification and regression tasks. For classification, this algorithm defines a hyperplane in a higher dimensional space to separate the features. The main goal of the algorithm is to find a plane with the highest margin (i.e., the maximum distance) between points for the different classes. Then, any new sample can be mapped into the same higher dimensional space and consequently assigned to a category based on the position of the corresponding points with respect to the planes that separate the different classes. SVM is mainly characterized by the *kernel* type considered and some hyper-parameters such as  $C$  and  $\gamma$ . For the sake of completeness, we recall that we call  $C$  the regularization parameter that controls the trade-off for achieving a low training error and a low testing error, and it determines the ability to generalize for the associated classifier with respect to unseen data. Higher values of  $C$  create a strict separation, but they may return a hyperplane that does not possess a large minimum margin between the two classes [6]. Instead, the parameter  $\gamma$  influences the curvature we want in the decision boundary whenever we make use of a polynomial or a sigmoid kernel.

To clarify the way  $C$  and  $\gamma$  work, let us consider training vectors  $x_i \in \mathbb{R}^p, i = 1, 2, \dots, m$ , divided in two classes, and a vector  $y = \{1, -1\}^m$ . Then, we aim to solve the following optimization problem [1]:

$$\min_{w, b, \zeta} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \zeta_i \quad \text{subject to } y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i, \text{ with } \zeta_i \geq 0, i = 1, 2, \dots, m \quad (1)$$

where  $b \in \mathbb{R}$ ,  $w$  stands for the normal vector to the hyperplane,  $\zeta_i$  is a slack variable designed to indicate the distance between the  $i^{th}$  voxel vector and the hyperplane,  $m$  is the amount of subjects in the dataset and  $C$  is the hyper-parameter previously mentioned. When we introduce a non-linear degree of separation between two classes, we generally make use of non-linear kernels when applying SVM, exploiting variations in the  $\gamma$  values [1]. In particular, we may have:

- Sigmoid kernel, with the following mathematical form:

$$\tanh(\gamma x_i^T x_j + r) \quad (2)$$

- Polynomial kernel, with the following mathematical form:

$$(\gamma x_i^T x_j + r)^d \quad (3)$$

with  $x_i$  being the  $i^{th}$  voxel vector,  $x_j$  representing the vectorized hyperplane,  $r$  being a constant and  $d$  the number of subjects in the system. Similarly to  $C$ , high values of  $\gamma$  may bring overfitting to the model.

RF is another supervised ML algorithm that combines several decision trees in order to obtain a single optimal result. A decision tree can be represented as a flowchart where internal nodes test the attributes of the data-samples, branches are outcomes of this test and terminal nodes (leaves) are labels. The choice of several hyper-parameters can influence the final result: some of the most critical are the number of decision trees included in the forest algorithm, the maximum number of features that should be taken into account to have a decision split and the highest number of levels that should be used in the algorithm training. When working in high-dimensional spaces, there are several drawbacks, such as the sparseness of the raw data and the computational intractability of the analysis; the highest amount of features that should be considered while making a decision split is particularly useful when the amount of features is high, since it can reduce the dimensionality of a mathematical problem. Furthermore, the selection of the minimum number of samples required to have a decision split or a leaf can be crucial [16].

### 2.3. Pre-processing operations

Some pre-processing operations have been taken into account before applying the ML algorithms above mentioned. First, we remark that the 3D images composing the dataset represent human brains without the skull structure, surrounded by a large area of empty voxels. Since such an area is of course useless for classification purposes, then it can be safely removed from each image. As a matter of fact, the useless region for each image has been first identified, and then the intersection between them has been extracted. Finally, just the intersection has been discarded from each image, this way preserving the same size for each sample of the dataset.

Then, in order to guarantee the same importance for all the features (voxels), the dataset has been standardized. In our experiments, we have tested three different scalers:

- Standard Scaler (SS) [22]: let  $X$  be the matrix containing all the data-samples in the columns and the features in the rows, while  $\mu$  and  $\sigma$  be the mean and standard deviation associated to each feature, respectively. Then, the standard scaler removes the mean and scales all the data to unit variance, namely:

$$X' = \frac{X - \mu}{\sigma} \quad (4)$$

- Robust Scaler (RS) [21]: let  $m$  be the median observation found within the dataset for each feature, while  $p_{75}$  and  $p_{25}$  be the 75-th and 25-th percentiles, respectively. Then the robust scaler removes the median and uses a scaling method relying on the interquartile range, such as:

$$X' = \frac{X - m}{p_{75} - p_{25}} \quad (5)$$

- Min-Max Scaler (MMS) [20]: let  $X_{\min}$  and  $X_{\max}$  be the minimum and maximum values for each feature in the dataset, respectively. Then, the Min-Max scaler scales the dataset between 0 and 1 by:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (6)$$

Furthermore, the use of Principal Component Analysis (PCA) [10] has been assessed to evaluate the performance impact of a dimensionality reduction in the system. PCA can be used to highlight primary patterns from a large set of data with the isolation of the most relevant features (or “components”) in the input information. However, since the use of PCA has provided more unstable results in our experiments, all the accuracy values included in section 3 have been collected without using PCA.

### 2.4. Description of the experiments

In order to find the best values for the hyper-parameters of both the algorithms, a grid search hyper-parameter optimization is computed at the beginning of our experiments. For SVM, we have considered  $C = \{0.1, 1, 10, 100, 1000\}$ ,  $\gamma = \{1, 0.1, 0.01, 0.001, 0.0001\}$ , while linear, polynomial and sigmoid functions have been tested as kernel.

For RF, we have analyzed the impact on the algorithmic training process by varying the number of decision trees between 200 and 2000 at a 10-step interval, the maximum depth of the trees between 10 and 109 levels at an 11-step interval, the minimum number of samples required to have a decision split ( $\{2, 5, 10\}$ ) and a leaf ( $\{1, 2, 4\}$ ).

For both SVM and RF algorithms, the grid-search technique is based on a 5-fold cross-validation that used accuracy as reference metric.

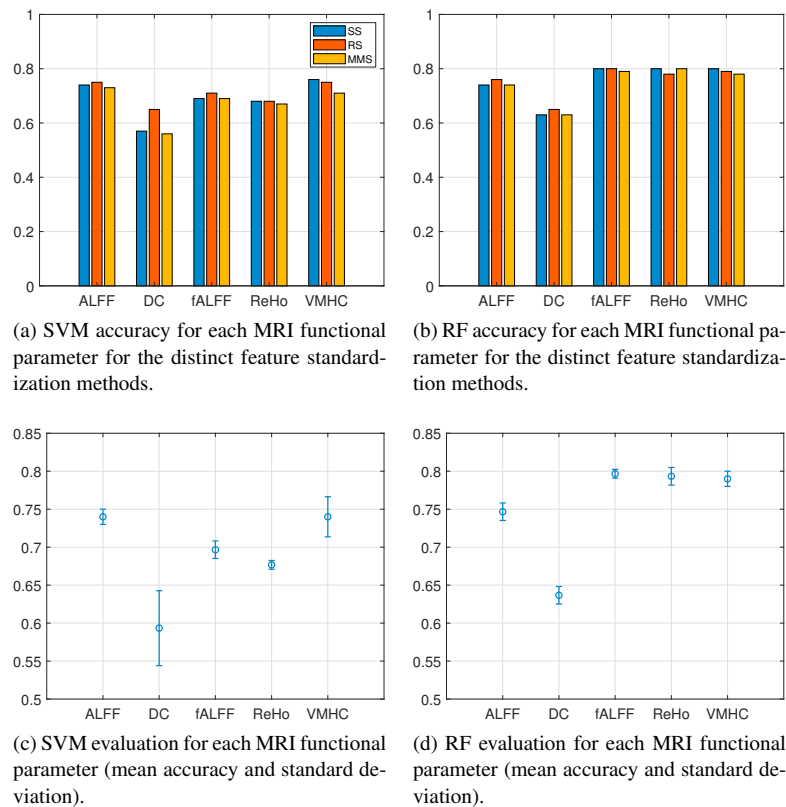


Fig. 2: Performance of SVM and RF in terms of accuracy using SSS.

Once the optimal hyper-parameters have been found, a Stratified Shuffle Split (SSS) strategy with 100 splits has been adopted for validation: it randomly splits data into train and test sets, by preserving the proportion of distribution of the information between train and test data. In this way, we allow to have the same amount of patients and controls in the resulting dataset, for all the 100 rounds of split.

As a detail, we remark that our experiments have been implemented using the Python programming language, with the functionalities provided by the scikit-learn library. The code is publicly available at [17].

### 3. Results and discussion

As a premise, we should state that our work primarily tested ML techniques. In order to avoid creating confusion to the eyes of the reader due to a series of statistically insignificant results, this section will exclusively focus on the SVM and RF techniques, where our research provided more insight on automated diagnostics.

The results obtained through the execution of the SVM algorithm provide an interesting insight into the information contained within the voxel of fMRI images. Fig. 2a reports the accuracy performance for each MRI functional parameter, showing the results of the different feature standardization methods considered in our experiments. In order to better appreciate the variability across the standardization methods, Fig. 2c indicates the associated mean and standard deviation accuracy. First, it is possible to observe consistent results among the different standardization techniques, with minor changes among methods. The average accuracy of the SVM methodology on the whole dataset on all the functional parameters and scalars is equal to 69%, at a standard deviation of 6%.

Looking more closely into the results obtained through the SVM technique, the ALFF and VMHC strategies achieve the best performance at 74% of mean accuracy, with a maximum accuracy recorded on robust (75% on ALFF) and standard (76% on VMHC) scalars. To give more context to the results, the classification accuracy obtained on the



other MRI functional parameters does not move above 71%. Interestingly, the result obtained on fALFF (71% with the robust scaler, with an average accuracy on all scalers of 70% and 1% standard deviation) is worse than that recorded on ALFF, indicating that the search for the optimal hyperplane to separate patients and controls in the dataset appears to be made more complex by the exclusion of the typical noise connected to mental states such as hunger from the MR images. Interestingly, this trend is recorded on all the three data scalers employed.

Furthermore, the DC technique reveals the worst results, a factor that may indicate the low relevance of single nodes composing the whole-brain network to spot the presence of FTD in elderly patients. This conclusion is confirmed by a coherent set of results recorded using different data scalers.

In terms of standard deviation, the results obtained on the ReHo technique provide the highest stability in the experiment, with a value of 0.6% and a mean accuracy rate measured at 68% (a value recorded both on average among the different data scalers and looking at the robust scaling technique, which provided the best result on ReHo). When applied to the more accurate results obtained on ALFF and VMHC, the same approach returns a standard deviation respectively equal to 1% and 3%, leading us to conclude that the SVM algorithm provides more reliable results on the ALFF technique.

By looking at the performances on different scalers, the robust methodology outperforms the standard and min-max techniques, having an average accuracy across all the MRI functional parameters equal to 71%, against, respectively, 69% and 68%. This result is certainly interesting, as it translates into a higher efficiency of the SVM algorithm when the dataset is elaborated to handle outliers in the data. The strategy to introduce a degree of distortion in the dataset by locating all the voxels in a space with mean equal to 0 and a fixed standard deviation proves to be less effective than the robust scaling. The same is true for the attempt performed by the min-max technique to scale the data while preserving their original structure.

The use of an RF methodology provides overall better results with respect to SVM across all the techniques and scaling methodologies, with an average accuracy on the whole dataset on all the pre-processing techniques and scalers equal to 75% and a slight increase in the standard deviation of the outcomes, which passes from 6% to 7% (see Fig. 2b and 2d). In order to compare the results obtained with the two algorithms, the experiments run with the RF strategy are validated with the same SSS validation function used on the SVM algorithm.

While ALFF and VMHC prove to have a valid informational value by using an RF algorithm, other techniques manage to lead the classifier to an even better output. In general terms, an improvement is recorded across all the MRI functional parameters and on each scaling methodology.

Furthermore, the creation of a RF of decision trees proves to find an actual improvement in the passage from ALFF to fALFF data elaboration. Specifically, the first technique leads the classifier to record a 74% accuracy as average value across all the scalers (peaking at 76% on the robust strategy), while the additional data normalization layer brought by fALFF allows the RF reach an 80% average accuracy. In both cases, the standard deviation is recorded at a low level, leading us to validate the interpretation of a prediction improvement. Actually, the standard deviation decreases following mean accuracy enhancement, moving from 1% to 0.5%.

An interesting confirmation of the results obtained through SVM comes from the evaluation on the DC functional parameter. While the mean accuracy improves (reaching 64% at 1% standard deviation, with robust scaling recording a 65% accuracy), the classifier proves to encounter significantly more complexities (meaning a decreased accuracy) while looking for the correct class discrimination on these images, compared to the other functional parameters. Therefore, the conclusion on the seemingly low importance of studying the single nodes composing the whole-brain network appears to be confirmed by the evidence on the RF algorithm.

As previously mentioned, VMHC returns positive results on the RF framework too, reaching a 79% average accuracy at 1% standard deviation, one of the limited cases in which the standard scaler outperforms the robust approach (with accuracies equal to, respectively, 80% and 79%). The same can be said about ReHo images (where the standard method records an 80% accuracy, against the 78% value obtained with the robust scaler), with the observation that, while this technique appears to have a below-average predictive power on SVM, the RF algorithm obtains a 79% average accuracy (with 1% standard deviation) on this subset of information.

Finally, the RF classifier confirms the evidence obtained from the SVM algorithm on the different scalers, with the robust technique recording a 76% accuracy against lower results obtained by the robust and min-max scalers (both amounting to 75%). We summarize all the results also in Tab. 1.

Table 1: Recap of SVM (top) and RF (bottom) accuracy for each MRI functional parameter for the distinct feature standardization methods. The mean and standard deviation are also reported.

	ALFF	DC	fALFF	ReHo	VMHC
Standard scaler	0.740	0.570	0.690	0.680	0.760
	0.740	0.630	<b>0.800</b>	<b>0.800</b>	<b>0.800</b>
Robust Scaler	0.750	0.650	0.710	0.680	0.750
	0.760	0.650	<b>0.800</b>	0.780	0.790
Min-Max Scaler	0.730	0.560	0.690	0.670	0.710
	0.740	0.630	0.790	<b>0.800</b>	0.780
Mean	0.740	0.593	0.697	0.677	0.740
	0.747	0.637	0.797	0.793	0.790
Standard deviation	0.010	0.049	0.012	0.006	0.026
	0.012	0.012	0.006	0.012	0.010

For the sake of completeness, we indicate the best hyper-parameters on the RF algorithm while using VMHC data and a Standard Scaler approach, in a framework where we obtained the most stable results:

- Number of trees = 1600
- Depth of the tree = 40
- Min sample for decision split = 2
- Min sample for decision leaf = 1

### 3.1. Further evaluation

We have extended our evaluation, by comparing the results of this work with those obtained by Premi et al. [18]. While the goals of that paper are not perfectly comparable to those of our research, a certain degree of similarity can be identified between the two sets of results. More precisely, Premi et al. studied only SVM as classification method, however they have not exploited any data scaling methodology. Furthermore, the  $C$  parameter of the algorithm has been manually tested in their work, not considering any hyper-parameter optimization.

Note that Premi et al. have implemented a Leave-one-out cross-validation (LOOCV), with  $n$  runs (where  $n$  is equal to the number of subjects in the dataset) having a single sample (singleton) as test, and the remaining data used to train the SVM. On the contrary, we have decided to exploit the advantages of SSS as a validation system instead of the LOOCV. Indeed, SSS works in such a way that the dataset is randomized and can be split into training and test subsets in a balanced way even thousands of times, leading the overall accuracy to possess a potentially larger statistical significance. This approach should provide more reliable results, considering that LOOCV is only composed of 30 (i.e., the number of subjects in the dataset) rounds of unbalanced split and validation.

However, to keep the comparison as fair as possible, we have evaluated our methods by considering LOOCV as validation. Fig. 3 and Tab. 2 report the results for each MRI functional parameter. Our accuracy refers to the RF algorithm and the best scaler for each technique. Remarkably, our highest accuracy has been obtained for VMHC technique (87%), and it outperforms Premi's best performance (83%). For the sake of completeness, we remark that in [18] the authors have applied the PCA technique to the data, since the corresponding dimensionality reduction has brought to accuracy improvements in their model. On the contrary, our experimental results show that training our models without PCA leads to more stable and discriminative results with respect to the use of the PCA.

### 3.2. Further experiments

To investigate the possibility of increasing the accuracy of our methods, we have tried to concatenate ALFF, fALFF, DC, ReHo and VMHC into one 3D array. This idea was to exploit the entire information contained in all the MRI



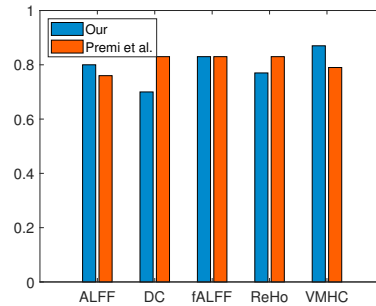


Fig. 3: Comparison with the work by Premi et al. [18] using LOOCV.

Table 2: Comparison with the work by Premi et al. [18] using LOOCV. In bold the highest accuracy returned by our method with VMHC.

	ALFF	DC	fALFF	ReHo	VMHC
Our	0.80	0.70	0.83	0.77	<b>0.87</b>
Premi et al.	0.76	0.83	0.83	0.83	0.79

functional parameters, instead of performing a separate analysis. However, in this case the highest accuracy (76%) has been reached using the SVM algorithm with Standard Scaler, proving a decay in the results obtained with respect to the use of a single MRI technique. Interestingly, this result indicates that integrating the information contained in the different functional parameters of MRI leads to a less discriminative feature space with respect to the feature space associated to the distinct MRI pre-elaborations.

Furthermore, the same experiments described in Sec. 2 have been performed on another dataset of 34 subjects: 17 asymptomatic carriers of GRN Thr272fs mutation and 17 young non-carriers belonging to the same families. Also this dataset has been acquired by the Centre of Aging and Neurodegenerative Diseases, and similarly divided into the ALFF, fALFF, DC, ReHo and VMHC functional parameters. However, the accuracy for both SVM and RF algorithms has been proved to be consistently close to 50%, demonstrating that it is currently impossible to distinguish the two classes of MRI given the current data available.

#### 4. Conclusion

The MRI classification between healthy and FTD-diagnosed subjects is a promising technique that can help experts in providing an early diagnosis which could be beneficial for the life quality of a patient. In this paper, we explored the performance of SVM and RF to formulate a correct FTD diagnosis. The dataset consisted of MRI data acquired on 30 subjects, and each set of data was elaborated with different MRI techniques based on a MVPA approach. The results obtained have been validated by a SSS process and RF achieved the best accuracy. When compared our performance with those of Premi et al., an improvement by 4% was obtained when using the VMHC technique. Importantly, the results showed a high consistency across the investigated methodologies, thus substantiating their reliability, even if the size of the dataset is relatively small. Of course, an increase in the original size of the dataset may lead to better results in terms of overall test accuracy, as well as to promote the exploration of Deep Learning algorithmic techniques.

#### Acknowledgements

We would like to show our gratitude to Enrico Premi and the other authors of [18] for sharing their wisdom and data on the topic. We are grateful both towards the Centre for Ageing Brain and Neurodegenerative Disorders (University of Brescia, Italy) and the colleagues working at the University of Turin (Italy) for all the insight and clarification over the methodology followed in their work.

## References

- [1] Amami, R., Ayed, D.B., Ellouze, N., 2015. Practical selection of SVM supervised parameters with different feature representations for vowel recognition. arXiv preprint arXiv:1507.06020 .
- [2] Arosio, B., Abbate, C., Galimberti, D., Rossi, P.D., Inglese, S., Fenoglio, C., Ridolfi, E., Gussago, C., Casati, M., Tedone, E., et al., 2013. Grn thr272fs clinical heterogeneity: a case with atypical late onset presenting with a dementia with lewy bodies phenotype. *Journal of Alzheimer's Disease* 35, 669–674.
- [3] Buckner, R.L., Sepulcre, J., Talukdar, T., Krienen, F.M., Liu, H., Hedden, T., Andrews-Hanna, J.R., Sperling, R.A., Johnson, K.A., 2009. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *Journal of neuroscience* 29, 1860–1873.
- [4] Chagué, P., Marro, B., Fadili, S., Houot, M., Morin, A., Samper-González, J., Beunon, P., Arrivé, L., Dormont, D., Dubois, B., et al., 2021. Radiological classification of dementia from anatomical MRI assisted by machine learning-derived maps. *Journal of Neuroradiology* 48, 412–418.
- [5] Chassagnon, G., Vakalopoulou, M., Paragios, N., Revel, M.P., 2020. Artificial intelligence applications for thoracic imaging. *European journal of radiology* 123, 108774.
- [6] Cristianini, N., Shawe-Taylor, J., et al., 2000. An introduction to support vector machines and other kernel-based learning methods. Cambridge university press.
- [7] Hashimoto, D.A., Witkowski, E., Gao, L., Meireles, O., Rosman, G., 2020. Artificial intelligence in anesthesiology: current techniques, clinical applications, and limitations. *Anesthesiology* 132, 379–394.
- [8] Hessler, G., Baringhaus, K.H., 2018. Artificial intelligence in drug design. *Molecules* 23, 2520.
- [9] Johnson, K.W., Torres Soto, J., Glicksberg, B.S., Shameer, K., Miotto, R., Ali, M., Ashley, E., Dudley, J.T., 2018. Artificial intelligence in cardiology. *Journal of the American College of Cardiology* 71, 2668–2679.
- [10] Jolliffe, I.T., 2002. Springer series in statistics. Principal component analysis 29.
- [11] Kim, J.P., Kim, J., Park, Y.H., Park, S.B., San Lee, J., Yoo, S., Kim, E.J., Kim, H.J., Na, D.L., Brown, J.A., et al., 2019. Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. *NeuroImage: Clinical* 23, 101811.
- [12] Knopman, D.S., Roberts, R.O., 2011. Estimating the number of persons with frontotemporal lobar degeneration in the US population. *Journal of Molecular Neuroscience* 45, 330–335.
- [13] Le, E., Wang, Y., Huang, Y., Hickman, S., Gilbert, F., 2019. Artificial intelligence in breast imaging. *Clinical radiology* 74, 357–366.
- [14] Niel, O., Bastard, P., 2019. Artificial intelligence in nephrology: core concepts, clinical applications, and perspectives. *American Journal of Kidney Diseases* 74, 803–810.
- [15] Olney, N.T., Spina, S., Miller, B.L., 2017. Frontotemporal dementia. *Neurologic clinics* 35, 339–374.
- [16] Pavlov, Y.L., 2019. Random forests. De Gruyter.
- [17] Pilipenko, T., Gnutti, A., Silvestri, A., Serina, I., Leonardi, R., 2022. ML Techniques for MRI Feature-Based Detection of FTD. <https://github.com/PilipenkoTatiana/ML-Techniques-for-MRI-Feature-Based-Detection-of-FTD>.
- [18] Premi, E., Cauda, F., Costa, T., Diano, M., Gazzina, S., Gualeni, V., Alberici, A., Archetti, S., Magoni, M., Gasparotti, R., et al., 2016. Looking for neuroimaging markers in frontotemporal lobar degeneration clinical trials: a multi-voxel pattern analysis study in granulin disease. *Journal of Alzheimer's Disease* 51, 249–262.
- [19] Rakić, M., Cabezas, M., Kushibar, K., Oliver, A., Lladó, X., 2020. Improving the detection of autism spectrum disorder by combining structural and functional MRI information. *NeuroImage: Clinical* 25, 102181.
- [20] Scikit-learn documentation, a. Min-Max Scaler. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html>.
- [21] Scikit-learn documentation, b. Robust Scaler. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html>.
- [22] Scikit-learn documentation, c. Standard Scaler. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>.
- [23] Xu, J., Zhang, M., 2019. Use of magnetic resonance imaging and artificial intelligence in studies of diagnosis of Parkinson's disease. *ACS chemical neuroscience* 10, 2658–2667.
- [24] Zhang, Z., Liu, Y., Jiang, T., Zhou, B., An, N., Dai, H., Wang, P., Niu, Y., Wang, L., Zhang, X., 2012. Altered spontaneous activity in Alzheimer's disease and mild cognitive impairment revealed by regional homogeneity. *Neuroimage* 59, 1429–1440.
- [25] Zou, Q.H., Zhu, C.Z., Yang, Y., Zuo, X.N., Long, X.Y., Cao, Q.J., Wang, Y.F., Zang, Y.F., 2008. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *Journal of neuroscience methods* 172, 137–141.
- [26] Zuo, X.N., Kelly, C., Di Martino, A., Mennes, M., Margulies, D.S., Bangaru, S., Grzadzinski, R., Evans, A.C., Zang, Y.F., Castellanos, F.X., et al., 2010. Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy. *Journal of Neuroscience* 30, 15034–15043.