Computerized Medical Imaging and Graphics 104 (2023) 102187

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

MRI texture-based radiomics analysis for the identification of altered functional networks in alcoholic patients and animal models

Silvia Ruiz-España [a], Rafael Ortiz-Ramón [b], Úrsula Pérez-Ramírez [a], Antonio Díaz-Parra [a], Roberto Ciccocioppo [c], Patrick Bach [d], Sabine Vollstädt-Klein [d], Falk Kiefer [d], Wolfgang H. Sommer [d,e], Santiago Canals [f,*,1], David Moratal [a,*,2]

[a] Center for Biomaterials and Tissue Engineering, Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain
[b] GRID Research Group, Universidad Internacional de Valencia – VIU, Valencia, Spain
[c] School of Pharmacy, University of Camerino, Camerino, Italy
[d] Department of Addiction Medicine, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany
[e] Institute of Psychopharmacology, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany
[f] Instituto de Neurociencias, Consejo Superior de Investigaciones Científicas, Universidad Miguel Hernández, Campus de San Juan, 03550 Sant Joan d'Alacant, Spain

ARTICLE INFO

Keywords:
Alcohol use disorders
Brain MRI
Striatal network
Radiomics
Standardized radiomic features
Classification algorithms

ABSTRACT

Alcohol use disorder (AUD) is a complex condition representing a leading risk factor for death, disease and disability. Its high prevalence and severe health consequences make necessary a better understanding of the brain network alterations to improve diagnosis and treatment. The purpose of this study was to evaluate the potential of resting-state fMRI 3D texture features as a novel source of biomarkers to identify AUD brain network alterations following a radiomics approach. A longitudinal study was conducted in Marchigian Sardinian alcohol-preferring msP rats (N = 36) who underwent resting-state functional and structural MRI before and after 30 days of alcohol or water consumption. A cross-sectional human study was also conducted among 33 healthy controls and 35 AUD patients. The preprocessed functional data corresponding to control and alcohol conditions were used to perform a probabilistic independent component analysis, identifying seven independent components as resting-state networks. Forty-three radiomic features extracted from each network were compared using a Wilcoxon signed-rank test with Holm correction to identify the network most affected by alcohol consumption. Features extracted from this network were then used in the machine learning process, evaluating two feature selection methods and six predictive models within a nested cross-validation structure. The classification was evaluated by computing the area under the ROC curve. Images were quantized using different numbers of gray-levels to test their influence on the results. The influence of ageing, data preprocessing, and brain iron accumulation were also analyzed. The methodology was validated using structural scans. The striatal network in alcohol-exposed msP rats presented the most significant number of altered features. The radiomics approach supported this result achieving good classification performance in animals (AUC = 0.915 ± 0.100, with 12 features) and humans (AUC = 0.724 ± 0.117, with 9 features) using a random forest model. Using the structural scans, high accuracy was achieved with a multilayer perceptron in both species (animals: AUC > 0.95 with 2 features, humans: AUC > 0.82 with 18 features). The best results were obtained using a feature selection method based on the *p*-value. The proposed radiomics approach is able to identify AUD patients and alcohol-exposed rats with good accuracy, employing a subset of 3D features extracted from fMRI. Furthermore, it can help identify relevant networks in drug addiction.

* Corresponding authors.
E-mail addresses: scanals@umh.es (S. Canals), dmoratal@eln.upv.es (D. Moratal).
[1] 91.95.61
[2] 87.72.76

https://doi.org/10.1016/j.compmedimag.2023.102187
Received 20 April 2022; Received in revised form 28 November 2022; Accepted 9 January 2023
Available online 21 January 2023

1. Introduction

Excessive alcohol consumption represents a major public health concern, contributing to 3 million deaths per year, a 5.3% of all deaths globally. Overall, the harmful use of alcohol accounts for a large burden of disease and injury, being responsible for 5.1% of all disability-adjusted life years (World Health Organization, 2018). Alcohol use disorder (AUD) is a mental health condition related to alcohol abuse and alcohol dependence, and it is one of the leading risk factors for death, disease, and disability (Dupuy and Chanraud, 2016). AUDs have a high prevalence, with the highest rate in the European Union (14.8% and 3.5%, men and women) (World Health Organization, 2018).

Although significant efforts are being made trying to understand the causes and outcomes of AUD, a better understanding of the brain mechanisms and pathophysiology underlying this complex phenomenon is still necessary. Neuroimaging studies, mainly based on Magnetic Resonance Imaging (MRI), have repeatedly revealed structural and functional brain abnormalities induced by alcohol drinking. Recent studies have proposed resting-state functional Magnetic Resonance Imaging (rs-fMRI) as a powerful tool in the study of the dynamic course of alcoholism and the evaluation of AUD effects (Fede et al., 2019; Zhu et al., 2018). By measuring the low-frequency fluctuations of the bold-oxygenated-level-dependent signals, it is possible to obtain the resting-state networks (RSNs), brain regions with temporally coherent activity (Karahanoglu and Van De Ville, 2015).

Alterations in various RSNs, as well as in their functional connectivity (FC), have been detected in AUD patients, particularly in the executive control, default mode, salience, attention, visual, and reward networks (Bordier et al., 2021; Chanraud et al., 2011; Kamarajan et al., 2020; Müller-Oehring et al., 2015; Weiland et al., 2014; Zhu et al., 2018, 2017) and similarly in animal models (Perez-Ramirez et al., 2017; Scuppa et al., 2020). However, the results reported are mixed, and some of the studies differ regarding the FC. While there are analyses showing a decreased connectivity in networks such as default mode, executive control, salience, and reward, others suggest increased connectivity (Kohno et al., 2017; Zhu et al., 2018). Therefore, although these studies have undoubtedly provided valuable information related to brain FC alterations in alcoholism, the use of alternative techniques may offer new insights into identifying consistent measures as potential AUD biomarkers.

In the last few years, radiomics analysis has attracted significant research interest, becoming a valuable source of imaging biomarkers, notably in the field of clinical oncology (Avanzo et al., 2017; Liu et al., 2019). However, the scope of radiomics applications is continuously growing and currently includes successful applications in neurology and psychiatry (Feng and Ding, 2020; Park et al., 2020; Shu et al., 2021). This approach focuses on extracting quantitative features from medical images by means of advanced analysis combined with machine learning methods, or any other method from the field of artificial intelligence, to improve the decision-support process noninvasively (Gillies et al., 2016). Although it involves several steps like image acquisition, image segmentation and pre-processing, feature extraction, feature selection, and model building, the heart of the process is the extraction of features that describe the region of interest quantitatively. Such radiomic features are often categorized into shape/size features, histogram-based, and texture-based features that are obtained by quantifying statistical inter-relationships between voxels and gray-level patterns. Radiomic features can also be extracted from filtered or mathematical transformed images (Rizzo et al., 2018; van Timmeren et al., 2020).

Several studies have applied machine learning techniques using features extracted from rs-fMRI to discriminate individuals with AUD from healthy controls, predict alcohol dependence or predict alcohol use severity. But in all cases, features related to FC within and between resting-state networks were almost exclusively used (Fede et al., 2019; Kamarajan et al., 2020; Zhu et al., 2018). We hypothesize that standardized radiomic features extracted from rs-fMRI can be used to discriminate the occurrence of excessive levels of alcohol drinking in subjects and, therefore, identify relevant networks in alcohol addiction. This hypothesis is based on the radiomics assumption that both structural and functional images likely contain quantifiable information reflecting the underlying tissue pathophysiology (Rizzo et al., 2018).

Therefore, the purpose of this study was to investigate the capability of 3D radiomic features extracted from rs-fMRI to identify subjects with AUD following a radiomics approach, thereby identifying key brain networks in alcoholism. We explored the potential of this analysis as a source of imaging biomarkers using a multilevel strategy: (1) applying the methodology to rs-fMRI data acquired in a rat model of AUD before and after alcohol intake. This longitudinal design conveys statistical power to the preclinical analysis, avoids confounding factors commonly associated with AUD patient comorbidities, and facilitates a causal link between the potential alterations found by texture analysis to alcohol drinking; (2) applying the methodology to cross-sectional rs-fMRI data acquired in patients with AUD and healthy controls; (3) validating the methodology in a different imaging modality, specifically using structural MRI scans in both, animals and humans. For the animal study, we used the Marchigian-Sardinian alcohol-preferring (msP) rats (Ciccocioppo et al., 2006), an animal model of excessive alcohol consumption which has previously demonstrated excellent translational validity (De Santis et al., 2020, 2019). For the human study, we used MRI data from a recently published clinical trial in treatment seeking AUD patients and healthy subjects (Bach et al., 2021, 2020).

2. Materials and methods

A general diagram showing the major steps of the proposed radiomics methodology is presented in Fig. 1.

2.1. Animal study

All experiments conducted on animals were approved by the Institutional Animal Care and Use Committee at the Instituto de Neurociencias, Alicante, Spain, and were performed in accordance with the Spanish (law 32/2007) and European regulations (EU directive 86/609, EU decree 2001–486 and EU recommendation 2007/526/EC).

Experiments were carried out using two cohorts of male msP rats, 18 rats in each cohort (370–480 g), imported from the breeding facility at the School of Pharmacy, University of Camerino, Italy. All the animals were singly housed under controlled conditions (temperature: 22 $\pm$ 2 °C, mean $\pm$ SD; relative humidity: 55 $\pm$ 10%, mean $\pm$ SD; 12-hour light/dark cycle) and with access to food and water at all times. For this purpose, transparent polycarbonate cages with bedding were used, and a wooden stick and nesting material were provided as a form of enrichment.

A first longitudinal study was conducted on animals of cohort 1 ($n = 18$). In this experiment, the rats had free-choice access to two drinking bottles (replaced every 2–3 days), one containing water and the other a 10% (v/v) of ethanol in water. Liquid intake and animal weight were recorded each time the bottles were replaced. Rats underwent MRI before alcohol access (control condition) and after 4 weeks of alcohol consumption (alcohol condition).

A second longitudinal study was conducted on animals of cohort 2 ($n = 18$). These rats were housed under the same conditions but only with water access to evaluate whether age may be influencing the brain alterations detected. These rats underwent MRI at the same time points.

2.2. Human study

The study was conducted at the Central Institute of Mental Health in Mannheim, Germany. The experimental groups consisted of 35 males, recently detoxified, abstinent alcoholics (age 45.91 $\pm$ 9.05, abstinence days 21 $\pm$ 7, 260 $\pm$ 120 [g]/day of alcohol pretreatment) and 33 healthy male volunteers (age 41.78 $\pm$ 9.48) recruited within the ERA-

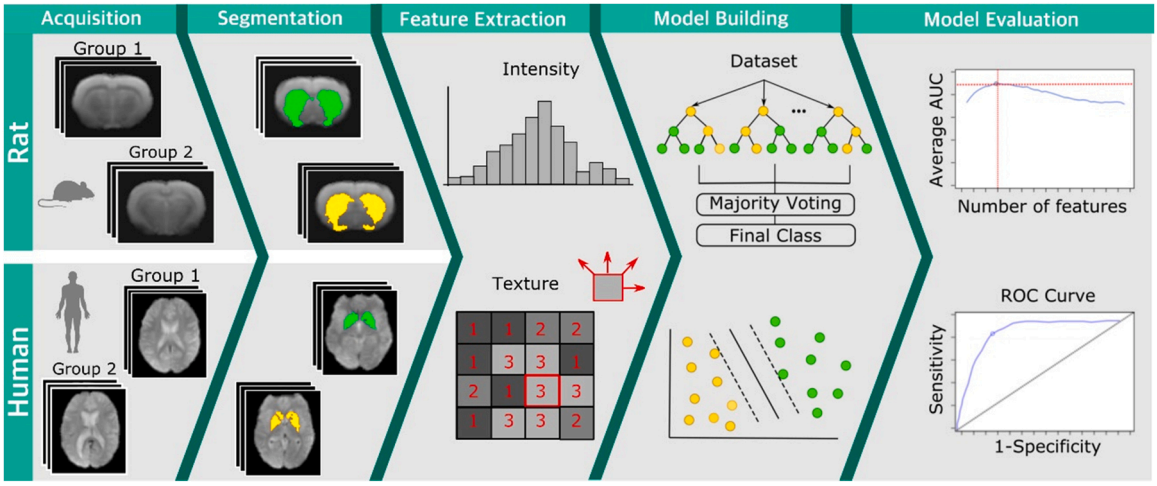

Fig. 1. Major steps of the proposed radiomics pipeline: image acquisition, region of interest definition (segmentation) and pre-processing, feature extraction, feature selection, and classification (model building and evaluation).

NET NEURON TRANSCALC study (WHO-International Clinical Trials Registry Platform: DRKS00003357). Analyses of resting-state fMRI, tasked-based fMRI, and diffusion data from this trial have been recently reported (Bach et al., 2021, 2020; Bordier et al., 2021; De Santis et al., 2020, 2019). Here we focus on the investigation of 3D radiomic features extracted from rs-fMRI data and hence included a subset of participants of the previously reported sample for whom high quality image data was available. Clinical characteristics are listed in the Supplementary Material (Table S1).

The key inclusion criteria for the AUD group was the diagnosis of alcohol dependence according to the Diagnostic Statistical Manual of Mental Disorders, DSM-IV (here equated to AUD), controlled abstinence of at least 2 weeks (21.91 ± 6.04 days, mean ± SD) before the MRI session, and completion of medically supervised detoxification (treatment of withdrawal symptoms with short acting benzodiazepines had to be completed for at least 3 days). Patients with psychiatric comorbidities or abuse of other substances (except smoking) were excluded. Details on the clinical assessment of patients are reported in the Supplementary Material, in the Extended Methods section. The local ethics committee approved the study procedures in accordance with the Declaration of Helsinki, and all participants gave written informed consent.

2.3. MRI acquisition protocols and image preprocessing: rat study

Imaging studies on rats were conducted on a 7 T, 30 cm bore size scanner (Bruker, Biospec 70/30, Ettlingen, Germany) equipped with a 675 mT/m and 11.4 cm inner diameter actively shielded gradient coil (Bruker, BGA 12-S). A 1 H rat brain receive-only phase array coil with integrated combiner and preamplifier, no tune/no match, was combined with an actively detuned transmit-only resonator (BrukerBioSpin MRI GmbH, Germany).

The same scanning parameters were used for all subjects since a variation in these parameters may affect the radiomic features and, as a consequence, the performance of the model (Rizzo et al., 2018; Waugh et al., 2011). Rs-fMRI acquisition was performed using a GE–EPI sequence with the following parameters: repetition time/echo time of 2000/15 ms; flip angle of 60º; matrix size of 96 × 96; pixel size of 0.26 × 0.26 mm^2; slice thickness of 1 mm and 90 time points consisting of 15 coronal slices. T2-weighted structural images were acquired using a rapid acquisition relaxation enhanced sequence (RARE) with the following parameters: repetition time/echo time of 2000/56 ms; RARE factor of 8; matrix size of 192 × 192; pixel size of 0.13 × 0.13 mm^2; slice thickness of 1 mm and 15 coronal slices.

Data were preprocessed using a combination of tools provided by the Oxford FMRIB software library (FSL; version 5.0, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki) (Jenkinson et al., 2012; Smith et al., 2004) and in-house scripts developed with MATLAB 2014a (The MathWorks, Inc., Natick, MA, USA, https://www.mathworks.com/). First, images were converted from Bruker to NIfTI (Neuroimaging Informatics Technology Initiative, https://nifti.nimh.nih.gov/) and voxel size resized by a factor of 10, a common step in rodents to fit human brain dimensions and apply the same algorithms (Kalthoff et al., 2011; Pan et al., 2015). Second, the fMRI data were motion corrected and segmented using the middle volume (middle time point) of each subject as reference (Jenkinson et al., 2002; Smith, 2002). Thereafter, two matrices were computed and concatenated to obtain the transformation matrix to be used in later steps to register the functional images to a rat brain template: a rigid matrix to co-register the functional images to the brain-extracted T2-weighted structural images and an affine matrix to normalize the structural images to the standard rat brain T2-weighted template (Schwarz et al., 2006). Afterward, spatial smoothing was applied for noise reduction using a Gaussian kernel with a 4 mm full-width at half-maximum (FWHM), and a global 4D mean-based intensity normalization was performed. A nuisance regression model was used to regress out head motion parameters (three translations plus three rotations and their derivatives) and linear trends. A band-pass temporal filtering was then applied (non-linear high-pass filter of $\sigma = 50$ s, and a Gaussian linear low-pass filter of $\sigma = 2$ s) to retain all frequencies ranging from 0.01 to 0.1 Hz (Pan et al., 2013). Finally, the filtered functional images were normalized to the standard space (Schwarz et al., 2006) by applying the computed transformation matrix (pixel dimensions: 0.19 × 0.19 × 0.8 mm^3).

For the analysis with the structural images, the inverse of the affine matrix was used to warp the regions under study from standard space to anatomical space and thus work with the native anatomical resolution.

2.4. MRI acquisition protocols and image preprocessing: human study

Imaging studies on humans were conducted on a 3 T Siemens MAGNETOM Trio TIM (Siemens AG, Munich, Germany) with a 32-channel head coil. Rs-fMRI acquisition was performed using a multiple gradient echo-planar T2 * -weighted pulse sequence with the following parameters: repetition time/echo time of 1500/28 ms; flip angle of 80º; matrix size of 61 × 73; pixel size of 3 × 3 mm^2; slice thickness of 3 mm and 230 time points consisting of 58 slices. Structural images were acquired using a T1-weighted magnetization prepared rapid acquisition gradient echo sequence (MPRAGE) with the following parameters: repetition time/echo time of 2300/3.03 ms; flip angle of 9º; matrix size

of 256 × 256; pixel size of 1 × 1 mm^2; slice thickness of 1 mm and 192 slices.

Data were preprocessed by combining the SPM8 software (Wellcome Trust Centre for Neuroimaging, London, UK, https://www.fil.ion.ucl.ac.uk/spm/software/spm8/) and in-house scripts developed with MATLAB. Firstly, data were corrected for physiological artifacts using Aztec (Van Buuren et al., 2009), thus removing the effects of heart rate and respiration acquired during scanning. Secondly, slice-timing correction was performed using the middle slice as reference (Sladky et al., 2011), images were also realigned to correct for head motion and co-registered with the structural images. After that, normalization to the standard MNI space (Montreal Neurological Institute, McGill University, Montreal, Canada) was performed, and spatial smoothing using a Gaussian kernel with a FWHM 8 mm was applied (Mikl et al., 2008). Finally, head motion parameters (three translations plus three rotations together with their derivatives) were regressed out, and a high-pass temporal filtering was applied (0.01 Hz).

For the analysis with the structural images and following the same procedure as in the animal study, the inverse of the matrix obtained in the normalization step was used to warp the regions under study from standard to anatomical space.

2.5. Region of interest definition and preprocessing

A fundamental question in radiomics analysis is where to locate the region of interest (ROI) to extract the most sensitive measurements. In this sense, using fMRI in radiomics studies could serve two functions. First, it allows the identification of data-driven regions of interest based on their activation dynamics in a resting state network analysis, potentially identifying disease-associated ROIs. Then, these functional ROIs can be used to focus the radiomics analysis since fMRI can also provide the voxel level data (BOLD contrast) for the texture analysis. In addition, using data-driven methods to define ROIs may lead to better predictions than using a predefined atlas (Dadi et al., 2019).

Therefore, a data-driven fMRI analysis method was used in this study to identify the regions of interest. To this end, a probabilistic independent component analysis (PICA) (Beckmann and Smith, 2004) was applied using MELODIC (FSL tool; version 3.14, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC).

The preprocessed data corresponding to control and alcohol conditions were temporally concatenated, whitened, and projected into a 17-dimensional subspace employing principal component analysis. The estimated component maps were divided by the standard deviation of the residual noise and thresholded to place an equal loss on false positives and false negatives (Beckmann and Smith, 2004). The obtained independent components (ICs) were visually inspected to identify those corresponding to RSNs and discard those associated with artifacts (Perez-Ramirez et al., 2017).

All ICs corresponding to RSNs were selected as ROIs, but ahead of feature extraction, some preprocessing was implemented within the ROIs to improve feature discrimination. Firstly, the 3D regions were resampled to the in-plane resolution using the cubic B-spline interpolation method. This step was performed only in those studies with non-isotropic voxels. Secondly, intensities were normalized between the range $\mu \pm 3\sigma$ (mean ± 3 SD of the gray level values) (Collewet et al., 2004). Finally, quantization to a smaller number of gray-levels (nGL) was performed to improve the signal-to-noise ratio and reduce the computation time (Gibbs and Turnbull, 2003). Four different nGL were analyzed and compared (16, 32, 64, and 128) to study their effect on the discriminatory power of the features.

2.6. Feature extraction

Forty-three features were extracted for each ROI using the Radiomics MATLAB package (Vallières et al., 2015) (https://github.com/mvallieres/radiomics). For the functional images, radiomic features were computed using the mean volume of the normalized data corresponding to the ROIs of each subject (Hassan et al., 2016). The extracted features follow the Image Biomarker Standardization Initiative (ISBI) (Zwanenburg et al., 2020), and they were computed based on five different statistical methods (summary in the Supplementary Material, Table S2): intensity histogram, gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size zone matrix (GLSZM) and neighborhood gray-tone difference matrix (NGTDM).

The texture features computed were rotationally invariant to avoid a possible source of bias in classification due to image orientation. To this end, in matrix-based methods, the neighboring properties of voxels in 13 directions of 3D space were averaged, obtaining only one matrix per method (Vallières et al., 2015). All feature values were standardized to zero mean and unit variance.

Afterward, a statistical analysis was performed across each selected ROI in the animal study. The main objective of this analysis was to evaluate the statistical significance of all the extracted features to determine if there are features that could be used as biomarkers to discriminate between control and alcohol conditions and thus identify the brain network with the most significant alterations in the pathological state. The features extracted from this brain network were then used for model building in both studies (animals and humans). For this purpose, features extracted in control and alcohol conditions across each ROI were compared using a Wilcoxon signed-rank test, and the *p*-values were adjusted for multiple comparisons using the Holm method. Alpha level was set at 0.05 for statistical significance.

2.7. Feature selection and classification

For the feature selection (FS) step, two different methods were employed to obtain a ranking of features ordered according to their relevance. First, a filter method based on the *p*-value was applied. This method scores each feature independently, without evaluating different combinations of features (Kuhn et al., 2013). To this end, the Wilcoxon signed-rank test was used in the animal study (paired samples) and the Mann-Whitney-Wilcoxon for independent groups of samples in the human study. This method was compared with a wrapper method, the support vector machine recursive feature elimination (SVM-RFE) technique (Guyon et al., 2002). To avoid overfitting, this step was implemented within the model-building process (Ambroise and McLachlan, 2002) and the models used the obtained rankings to select the optimal number of features.

Six predictive models were evaluated using the Caret package (Kuhn, 2008) implemented in R language, version 3.2.5 (R Development Core Team, Vienna, Austria, https://topepo.github.io/caret/). These models were selected to take into account different families of classifiers and because of their well-known performance in other datasets (Fernández-Delgado et al., 2014):

- Naïve Bayes (NB): a Gaussian kernel was used to estimate the probability density function.
- K-nearest neighbors (KNN): the Euclidean distance was used as the distance metric, and the optimum number of neighbors was chosen from {1, 3, 5, 7, 9, 11, 13, 15} in the parameter tuning process.
- Random forest (RF): the number of trees was set to 250, and the number of variables randomly sampled as candidates at each split was chosen from {2, 4, 6, 8, 10, 12, 14} in the parameter tuning process.
- Multilayer perceptron (MLP): one hidden layer was used, with several nodes chosen from {3, 5, 7, 9, 11, 13, 15} in the parameter tuning process.
- Support vector machine – linear (SVM_L): the SVM with linear kernel was used, and the cost parameter was chosen from $\{2^{-4}.2^{4}\}$ in the parameter tuning process.
- Support vector machine – radial (SVM_R): the SVM with Gaussian kernel was used. The cost and kernel spread parameters were chosen

from $\{2^{-4}.2^{4}\}$ and $\{10^{-2}.10^{2}\}$, respectively in the parameter tuning process.

Model performance was evaluated using a nested cross-validation (CV) structure (Supplementary Material, Fig. S1). A leave-group-out CV (LGOCV) was implemented in an outer loop to randomly split the dataset into training (75%) and testing (25%) a total of N = 100 times, forming N groups. In each group, the training set was used to obtain the ranking of features, and the classification accuracy was computed for all possible subsets of features. The feature subsets were generated by progressively adding the features one by one according to the ranking. Each feature subset was used to execute the parameter tuning process of the classification model (10-fold CV inner loop), so different parameter combinations were evaluated across the folds to select the optimal ones. The optimal parameters were then used to train a definitive model using the whole inner loop data and to compute the classification accuracy on the testing set in the outer loop. Classification performance was computed by averaging the area under the curve (AUC) of the receiver operating characteristic (ROC) over groups' estimates (mean ± SD). The optimal number of features was chosen for each model by comparing the classification accuracy for the different feature subsets. Sensitivity and specificity were also computed.

3. Results

3.1. Animals

During the four-week period of alcohol access, the animal average

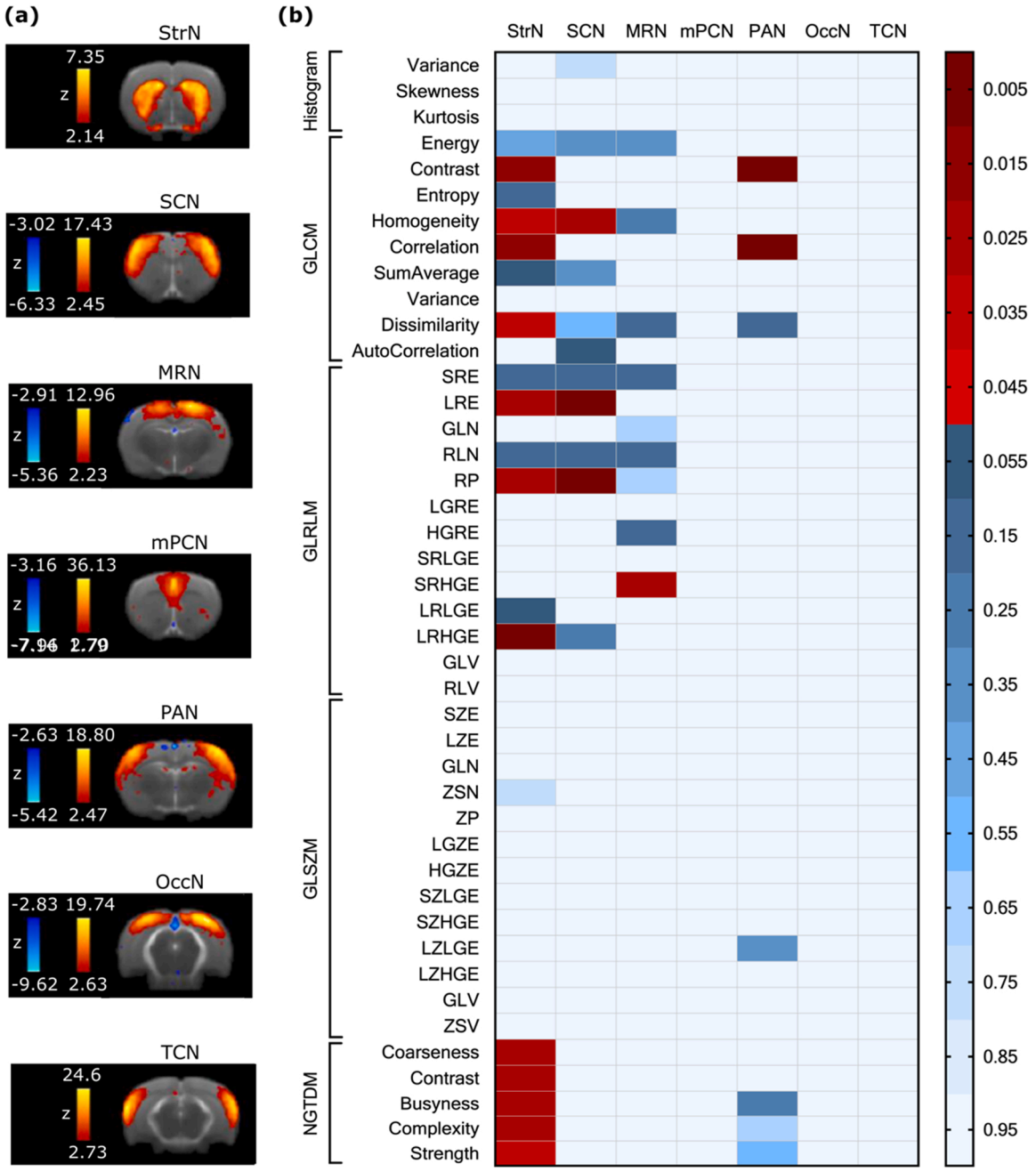

Fig. 2. (a) Resting-state networks obtained with group-PICA considering control and alcohol conditions (Perez-Ramirez et al., 2017). (b) Heatmap showing the results of the statistical analysis performed for nGL= 16 across each selected ROI. The intensity of color represents the Holm-corrected *p*-value. The red tones indicate statistically significant values.

alcohol intake was 5.8 ± 1.4 g/kg/day (mean ± SD).

3.1.1. Texture alterations in the different RSNs

Seven relevant RSNs were identified using the group-PICA analysis and selected as ROIs: striatal network (StrN), sensory-cortex network (SCN), motor-retrosplenial network (MRN), medial prefrontal-cingulate network (mPCN), parietal association network (PAN), occipital network (OccN), and temporal-cortex network (TCN). The most representative slice of each network is shown in Fig. 2(a).

We next compared all the extracted features in all ROIs before and after alcohol drinking to evaluate their statistical significance. Considering all quantization levels, forty-three features from four ROIs showed statistical significance ($p < 0.05$, Wilcoxon signed-rank test, multiple comparisons corrected by the Holm method), and twenty-three of them derived from the same ROI, the StrN. On the contrary, three RSNs did not provide any significant feature: mPCN, OccN, and TCN. Between six and eight significant features derived from each of the remaining three ROIs. It is noteworthy that most statistically significant features were obtained from regions quantized with a lower number of gray-levels, highlighting the results obtained with nGL= 16 (Fig. 2(b)).

Considering these first-stage results, the StrN seems to be the network better classifying the pathological state; hence features extracted from this network were selected for subsequent analysis.

3.1.2. Classification analysis using the striatal network

Longitudinal analysis in the StrN comparing naïve and alcohol conditions was performed with six different classifiers and two feature selection methods, achieving high classification accuracy (AUC > 0.85) in all cases.

The highest AUC value was obtained using features from the ROI quantized with nGL= 32 and the SVM-RFE feature selection method, achieving an AUC of 0.915 ± 0.100 with the top 12 ranked features. For this FS method, the best classification performance in terms of AUC for the six models is reported in Table 1.

As can be seen in Table 1, the best results for each model were obtained for different nGL. In fact, the nGL used in the quantization process affected the performance of the models differently (Fig. 3(a)). In general, features from the ROI quantized with nGL of 32 and 64 provided better classification accuracy, losing discriminative power specially when quantizing with nGL= 128. RF and SVM_R outperformed the rest of the models for all nGL. The optimal nGL in both cases was 32 with better performance with the RF model (paired Wilcoxon test, $p < 0.01$). Regarding the computational time, it is worth mentioning that using a computer with an Intel Core i5 processor of 2 GHz with 16 GB of RAM running on macOS Catalina (v. 10.15) operating system, the training time of the definitive RF model was 6 min, and 48 s. However, the training time of the SVM_R was almost triple (18 min and 58 s).

Table 1
Best classification results obtained for all models using the SVM-RFE feature selection method (animal study).

Model	nGL	# features	AUC*	Sensitivity*	Specificity*
NB	64	2	0.859 ± 0.155	0.832 ± 0.198	0.830 ± 0.173
KNN	64	39	0.872 ± 0.121	0.905 ± 0.149	0.740 ± 0.173
RF	32	12	**0.915 ± 0.100**	0.847 ± 0.200	0.822 ± 0.171
MLP	32	36	0.886 ± 0115	0.780 ± 0.204	0.78 ± 0.204
SVM_L	64	6	0.898 ± 0.146	0.800 ± 0.221	0.830 ± 0.173
SVM_R	32	9	0.906 ± 0.094	0.590 ± 0.335	0.855 ± 0.174

* mean ± SD as a result over groups' estimates.
Sensitivity and specificity were computed according to the optimal cutoff point of the ROC
curve, measured with the 'closest-to-(0,1)' criterion.
The highest AUC is highlighted in bold.
Abbreviations: NB – Naïve Bayes, KNN – K nearest neighbors, RF – Random Forest,
MLP – Multilayer perceptron, SVM_L – Support vector machine – linear, SVM_R – Support vector machine- radial, nGL – number of gray-levels, AUC – Area under the ROC
curve.

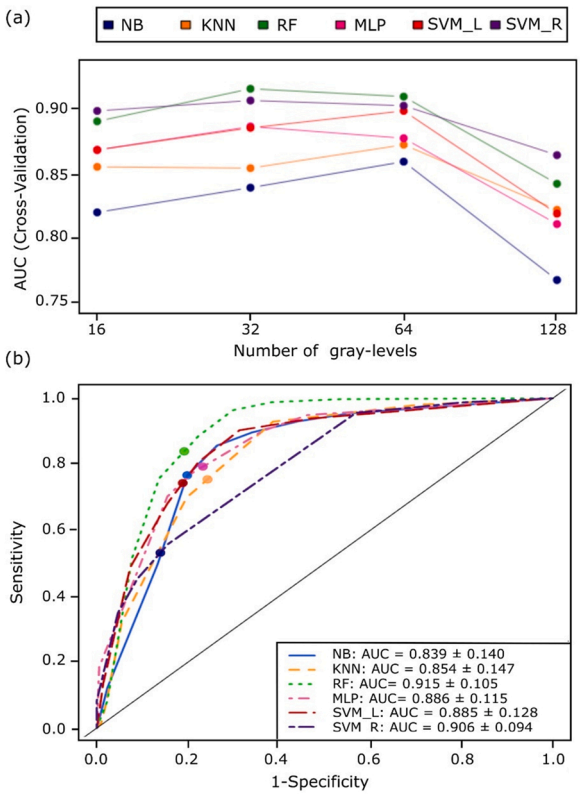

Fig. 3. (a) Comparison among the maximum AUC results obtained using features extracted from the StrN quantized with four different nGL and applying the six models under study with the SVM-RFE feature selection method. (b) Average ROC curves for the six models using the SVM-RFE feature selection method (nGL=32). The highlighted points on the curves indicate the optimal cutoff points that weight sensitivity and specificity equally, measured with the 'closest-to-(0,1)' criterion.

When using the model with the highest AUC (RF model) but with features extracted after quantizing with nGL= 64, the results were also good (AUC = 0.909 ± 0.114, 6 features), although significantly lower than those obtained with nGL= 32 (paired Wilcoxon test, $p < 0.01$). Representative ROC curves for the six models when using features extracted from the StrN quantized with nGL = 32 are shown in Fig. 3(b).

These results were compared with those obtained using the *p*-value FS method to check if ranking the features according to different criteria improved the classification results and reduced the optimal number of features (Fig. 4(a)). The best overall performance using the filter method was also achieved with the RF model (AUC = 0.911 ± 0.117, 24 features), without statistically significant differences but increasing the number of features used. In fact, the optimal number of features increased notably for all models, which reduces the chances of the models being generalized and increases the computational cost. Taking this into account, selecting the SVM-RFE method would be the best option in this case, as it uses less than half of the features to achieve similar AUC values.

The top 15 features returned by the SVM-RFE method averaged over cross-validation are shown in Fig. 4(b). This top 15 of the ranking included features from all feature extraction methods. However, most of

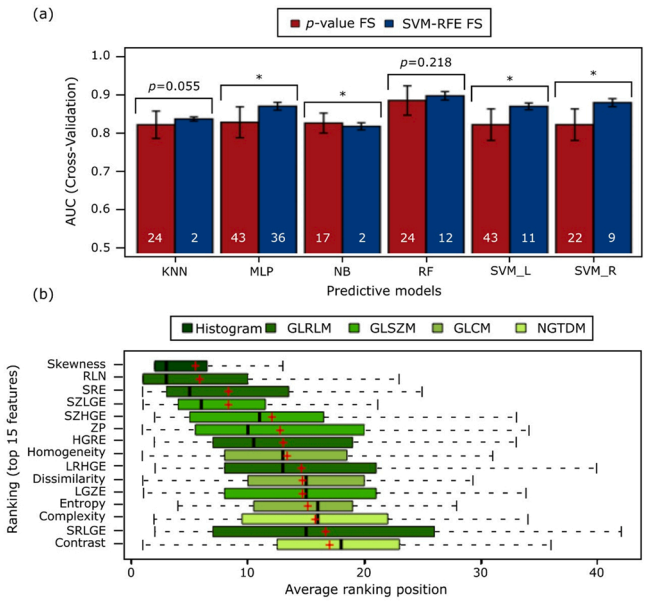

Fig. 4. (a) Barplot comparing the mean and standard deviation of AUC values obtained with the six models (nGL=32) using two FS methods: a filter (*p*-value) and a wrapper (SVM-RFE). The asterisks indicate a statistically significant difference in the classification performance (paired Wilcoxon test). The white numbers indicate the number of features used to achieve the maximum AUC. Error bars represent SD. (b) The top 15 features mainly contributing to discriminating between control and alcohol conditions using the SVM-RFE method. Features are classified by average position in the ranking, so lower values indicate higher importance of the feature. The red crosses indicate the average position, and the boxplots the distribution of importance over cross-validation. Centre lines within boxplots represent the median value, box limits indicate the 25th and 75th percentiles, and whiskers extend to the extreme values within 1.5 times the interquartile range from the upper and lower quartile.

the features used by the RF model (8 out of the 12) derived from the GLRLM and GLSZM methods.

3.1.3. *Influence of animal age*

Next, it was checked whether changes detected in the StrN network were solely due to alcohol drinking or had an ageing-related contribution. We tested on animals of cohort 2 whether 30 days without alcohol exposure (alcohol-naïve condition) induced any alteration in the network.

For this purpose, the model with the best performance in terms of AUC (RF model, nGL= 32, SVM-RFE FS method, 12 features) was selected. The model was trained in cohort 1 (control and alcohol conditions) using this best subset of features, and it was tested in cohort 2 (naïve-control conditions). Both cohorts were preprocessed in the same way, and the classification parameters and procedures remained identical. Applying this procedure, it was achieved an AUC = 0.5030, so the model was not able to detect changes between both time points in the animals of cohort 2. Since, in the previous section, the results obtained with the RF model using the *p*-value FS method (nG = 32, 24 features) were similar in terms of AUC (without statistical significance, $p = 0.218$), we repeated the procedure with this model, obtaining an AUC = 0.467. Therefore, we can conclude that ageing in the 1-month period between the two imaging time points had no significant contribution to the alterations detected in the striatal network.

3.1.4. *Influence of data preprocessing: global signal regression*

The inclusion of the global signal as a nuisance regressor in the functional MRI preprocessing is a step that generates a great deal of confusion and controversy (Liu et al., 2017). To check if this step produced significant changes in the feature outcome, the entire pipeline was repeated, modifying solely the data preprocessing step. In this case, the global signal was included as a nuisance regressor.

The results obtained with the models built using the data preprocessed with global signal regression were very similar to those obtained with the models without global signal regression. The highest AUC value was also obtained using the RF model and the SVM-RFE FS method, with nGL= 32 and 7 features, achieving an AUC of 0.912 ± 0.102. In fact, there were no statistically significant differences between these results and those obtained without signal regression for any of the models (Supplementary Material, Fig. S2).

Therefore, it could be verified that this preprocessing step did not produce significant differences in the feature outcome and, consequently, in the final classification results.

3.2. *AUD patients*

The same pipeline was applied to the human fMRI data set. The relevance of the striatal region in AUD, and that of the insular cortex, was recently highlighted in a network analysis using this dataset (Bordier et al., 2021). Therefore, for comparison with the animal model results, we selected the striatal region and a ROI equivalent to rat StrN, composed of the caudate putamen, nucleus accumbens, and ventral pallidum.

All features extracted from the selected region were analyzed using the six classifiers and both feature selection methods. The highest AUC values were obtained using the *p*-value FS method when quantizing the ROIs with 16 and 32 nGL (Table 2), decreasing the classification performance when increasing the number of gray levels, as found in the animal study (Fig. 5(a)). The results obtained when quantizing with nGL= 128 were very poor (AUC < 0.544).

As shown in Fig. 5(a), SVM_L and MLP models slightly outperformed the rest of the models in terms of AUC, with almost the same maximum AUC values for most nGL. The lower performance of all models with human data was expected due to the higher variability inherent to the clinical population and, most importantly, to the cross-sectional vs. longitudinal design in the human vs. the rat study, respectively. Therefore, with the human sample, we studied the distribution of the results in more detail to determine which was the model with the best

Table 2
Best classification results obtained for all models using the *p*-value feature selection method (human study).

Model	nGL	# features	AUC*	Sensitivity*	Specificity*
NB	16	2	0.710 ± 0.120	0.619 ± 0.157	0.716 ± 0.148
KNN	32	2	0.731 ± 0.124	0.690 ± 0.153	0.672 ± 0.162
RF	16	9	0.724 ± 0.117	0.646 ± 0.184	0.673 ± 0.164
MLP	16	2	0.745 ± 0123	0.672 ± 0.154	0.698 ± 0.168
SVM_L	16	2	**0.746 ± 0.123**	0.658 ± 0.157	0.743 ± 0.165
SVM_R	32	2	0.712 ± 0.116	0.583 ± 0.174	0.706 ± 0.175

* mean ± SD as a result over groups' estimates.
Sensitivity and specificity were computed according to the optimal cutoff point of the ROC
curve, measured with the 'closest-to-(0,1)' criterion.
The highest AUC is highlighted in bold.
Abbreviations: NB – Naïve Bayes, KNN – K nearest neighbors, RF – Random Forest,
MLP – Multilayer perceptron, SVM_L – Support vector machine – linear, SVM_R – Support vector machine- radial, nGL – number of gray-levels, AUC – Area under the ROC
curve.

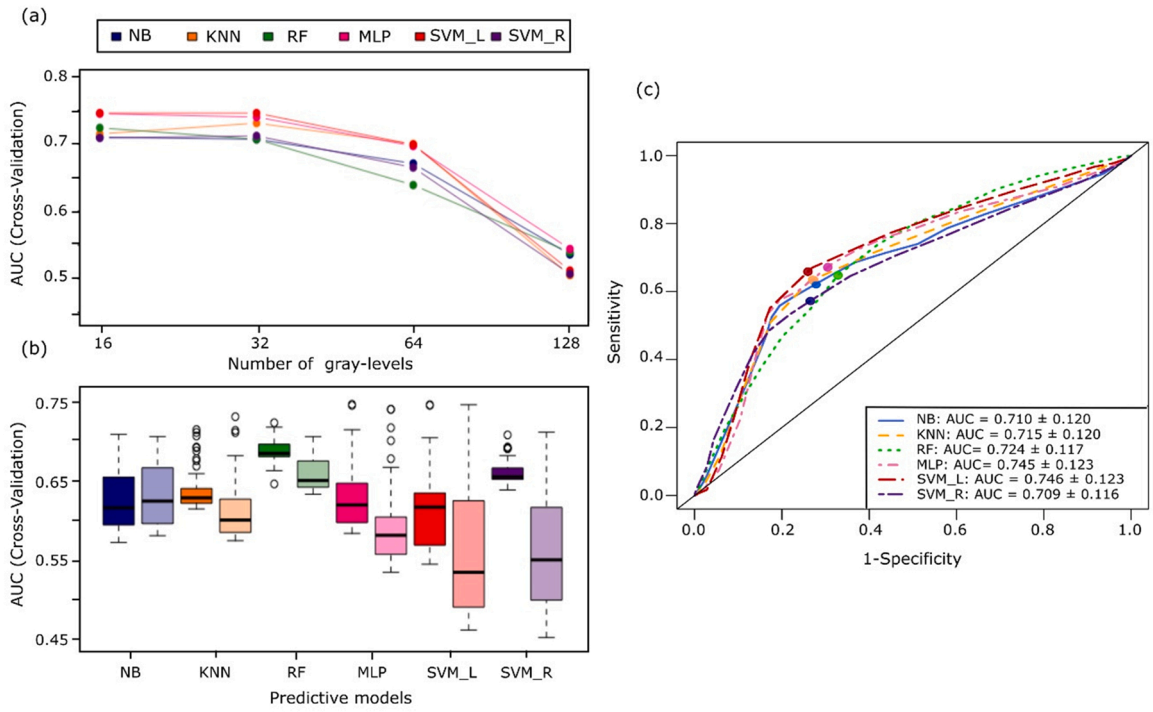

Fig. 5. (a) Comparison among the maximum AUC results obtained using features extracted from the ROI quantized with four different nGL and applying the six models under study with the *p*-value FS method. (b) Boxplots of the AUC results for the six models (*p*-value FS method) with nGL= 16 (dark color) and nGL= 32 (light color). (c) Average ROC curves for the six models using the *p*-value FS method (nGL=16). The highlighted points on the curves indicate the optimal cutoff points that weight sensitivity and specificity equally, measured with the 'closest-to-(0,1)' criterion.

classification performance (Fig. 5(b)). Interestingly, we observed that the RF was the model with the highest median and the least variability for both 16 and 32 nGL. In general, less variability was observed for most of the models in the results obtained with nGL = 16. The opposite occurred with the model with the highest AUC (SVM_L), being the model with the lowest median and highest variability in the results. Therefore, RF was again considered the model with the best overall performance, achieving a maximum AUC of 0.724 ± 0.117 with the top 9 ranked features after quantizing the images with nGL= 16, a similar value to the AUC obtained with the MLP and SVM_L models. ROC curves for the six models using features extracted from the ROI quantized with nGL = 16 are shown in Fig. 5(c).

These results were compared with those obtained using the SVM-RFE feature selection method (Fig. 6(a)). In this case, the best overall performance was also achieved with the RF model (AUC = 0.692 ± 0.120). Still, the optimal number of features was doubled, and the results were significantly lower than those obtained with the *p*-value FS method ($p < 0.01$). Maximum AUC values were higher using the *p*-value FS method for all the models, and the number of features used was lower.

The top 15 features returned by the *p*-value method averaged over cross-validation are shown in Fig. 6(b). Most of the features used by the RF model (7 out of the 9) derived from the GLRLM and GLSZM feature extraction methods, as also found in the animal model. However, in this case, it is worth highlighting two features derived from the GLCM, which always appeared in the ranking in the first two positions. These features are the Sum Average and the Autocorrelation.

3.3. Influence of iron accumulation

Previous evidence suggests an association between different neuropsychiatric disorders and abnormalities in brain iron content, including AUD, which is associated with excess iron accumulation. Brain iron concentration has been previously studied in AUD patients using rs-fMRI, and higher iron levels have been detected in deep grey matter regions such as the caudate nucleus, putamen, globus pallidum, and dentate nucleus (Juhás et al., 2017). An increase of brain iron has also been found in animals with acute and chronic alcohol exposure (Crews and Nixon, 2009; Rouach et al., 1990). Iron deposits in the brain result in a signal dropout in fMRI and other imaging modalities, being more significant the effect when the field strength increases (Pfefferbaum et al., 2010).

Therefore, we checked whether a change in iron metabolism could contribute to the texture results. To detect changes in iron accumulation, we compared the signal intensity in the T_2 * -weighted images of control and alcohol conditions in rats and humans. In the animal study, there was a slight signal loss in the alcohol condition but did not reach significance ($p = 0.1297$, paired Wilcoxon signed-rank test). Interestingly, in the human study, there was a slight increase in signal in the ROIs corresponding to AUD patients but also without statistically significant differences ($p = 0.1003$, Mann-Whitney-Wilcoxon test for independent groups of samples). The results of the analyses are shown in the Supplementary Material, Fig. S3. These preliminary results do not suggest differences in iron accumulation as a significant contribution to the alterations detected with texture analysis in the regions analyzed.

3.4. Structural MRI

Finally, the developed methodology was validated using a different imaging modality. We studied the capability of radiomic features to identify AUD subjects using structural MRI scans in animals and humans.

In the animal study, high classification (AUC > 0.93) was obtained for the six models (Table 3). In general, the best classification performance was obtained after quantizing the ROIs with nG = 16 and using the *p*-value FS method, decreasing the classification accuracy slightly when increasing the nGL (Fig. 7(a, b)). The highest AUC result was achieved with the MLP model followed by the SVM_R (nGL= 32), with no statistical significance between these results (paired Wilcoxon test,

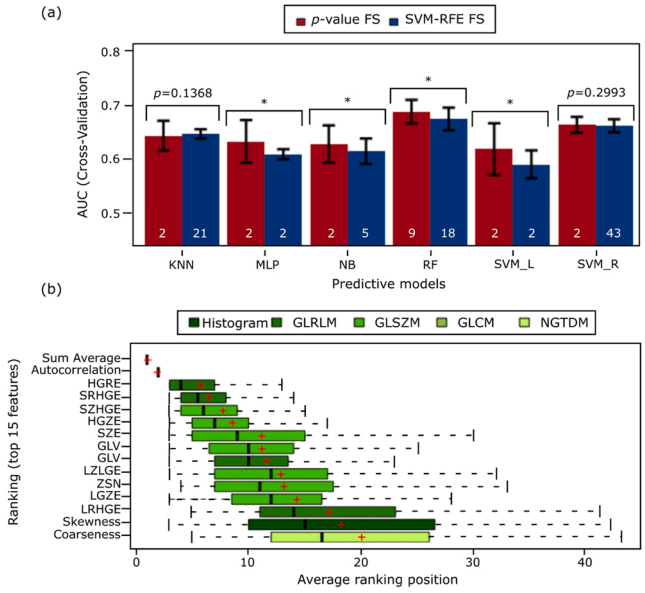

Fig. 6. (a) Barplot comparing the mean and standard deviation of AUC values obtained with the six models (nGL= 16) using two FS methods: a filter (*p*-value) and a wrapper (SVM-RFE). The asterisks indicate a statistically significant difference in the classification performance (paired Wilcoxon test). The white numbers indicate the number of features used to achieve the maximum AUC. Error bars represent SD. (b) The top 15 features mainly contributing to discriminating between control and alcohol conditions using the *p*-value FS method. Features are classified by average position in the ranking, so lower values indicate higher importance of the feature. The red crosses indicate the average position, and the boxplots the distribution of importance over cross-validation. Centre lines within boxplots represent the median value, box limits indicate the 25th and 75th percentiles, and whiskers extend to the extreme values within 1.5 times the interquartile range from the upper and lower quartile.

Table 3
Best classification results obtained for all models using the *p*-value feature selection method (structural scans).

	Animal study			Human study		
Model	**nGL**	**# features**	**AUC***	**nGL**	**# features**	**AUC***
NB	16	3	0.936 ± 0.092	16	6	0.746 ± 0.099
KNN	16	2	0.939 ± 0.112	64	9	0.761 ± 0.103
RF	16	42	0.944 ± 0.063	64	10	0.764 ± 0.094
MLP	16	2	**0.955 ± 0087**	64	18	0.821 ± 0.104
SVM_L	16	3	0.942 ± 0.097	16	43	**0.823 ± 0.098**
SVM_R	32	23	0.949 ± 0.067	16	43	0.768 ± 0.102

* mean ± SD as a result over groups' estimates.
Sensitivity and specificity were computed according to the optimal cutoff point of the ROC
curve, measured with the 'closest-to-(0,1)' criterion.
The highest AUC is highlighted in bold.
Abbreviations: NB – Naïve Bayes, KNN – K nearest neighbors, RF – Random Forest,
MLP – Multilayer perceptron, SVM_L – Support vector machine – linear, SVM_R – Support vector machine- radial, nGL – number of gray-levels, AUC – Area under the ROC
curve.

$p = 0.073$). However, with the MLP model, a considerably smaller number of features was necessary to achieve the best classification, thus facilitating its application to other datasets and considerably reducing the computational time. The training times of the definitive MLP and SVM_R models were 13 min and 11 s, and 29 min and 53 s, respectively. The best subset of features used by the MLP model (2 features) was derived from the GLRLM method. The top 15 features returned by the *p*-value method averaged over cross-validation are shown in Supplementary Material, Fig. S4 (a).

In the human study, good classification results were obtained after quantizing the images with nGL= 16 and nGL= 64, and using the *p*-value FS method (Table 3). MLP and SVM_L outperformed the rest of the models for all numbers of gray levels (Fig. 7(c)). Regarding these two models, the highest AUC value was achieved using the SVM_L, with results significantly higher (paired Wilcoxon test, $p < 0.01$) than those obtained with the same model and nGL= 64 (AUC = 0.821 ± 0.098). However, using this model, all features were necessary to achieve optimal classification performance (43 features). On the other hand, good AUC results were also achieved when using the MLP model (nGL = 64), and, in this case, the number of features used to obtain the highest AUC value was notably lower (18 features). Although both models got similar outcomes, models with fewer features are better candidates to be generalizable. ROC curves for the six models using features extracted from the ROI quantized with nGL = 64 are shown in Fig. 7(d). The best subset of features used by the MLP model included features from all the texture extraction methods. However, more than half of the features (10 texture features) were derived from the GLSZM and the GLRLM methods. The top 15 features returned by the *p*-value method averaged over cross-validation are shown in Supplementary Material, Fig. S4 (b).

4. Discussion

The focus of this translational study was two-fold. First, proposing a 3D texture analysis combined with machine learning techniques on rs-fMRI following a radiomics approach able to discriminate individuals with AUD and, second, identify key brain networks in alcohol addiction based on the texture analysis.

A first-stage analysis where seven RSN were identified in msP rats provided insight into possible brain targets in alcoholism, with the striatal network accumulating the largest statistically significant differences between image features after alcohol exposure. This finding was in agreement with recent human literature where brain alterations involving this network and associated with alcohol use have been detected in AUD patients (Kohno et al., 2017; Müller-Oehring et al., 2015; Perez-Ramirez et al., 2017). The proposed radiomics approach applied both to the msP rats and a cohort of AUD patients and matched healthy controls, supported these findings, showing that features extracted from the striatal region discriminated both alcohol conditions with good accuracy in rats (AUC > 0.9) and humans (AUC > 0.72).

Convergent results were found in both species, achieving the best overall performance using the Random Forest model in both cases. Previous studies also found that this model had higher prognostic performance than other predictive models (Rizzo et al., 2018), and it has shown the potential to detect adverse alcohol-related effects (Fede et al., 2019; Kamarajan et al., 2020; Zhu et al., 2018). In this work, the models obtaining the highest AUC values were not always considered the best models throughout the different analyses. In this sense, it is important to highlight the results obtained with the SVM model, achieving the highest accuracy on several occasions, such as in the human study using functional features. However, it was also one of the models with higher variability in the results. On the contrary, the RF model was generally the model with the highest median and least variability achieving similar AUC values, so it was considered the model with the best overall performance in the functional analysis. One of the advantages of the RF model is its relatively smaller bias and lower variance, increasing its generalization power (Kamarajan et al., 2020). Also, the number of

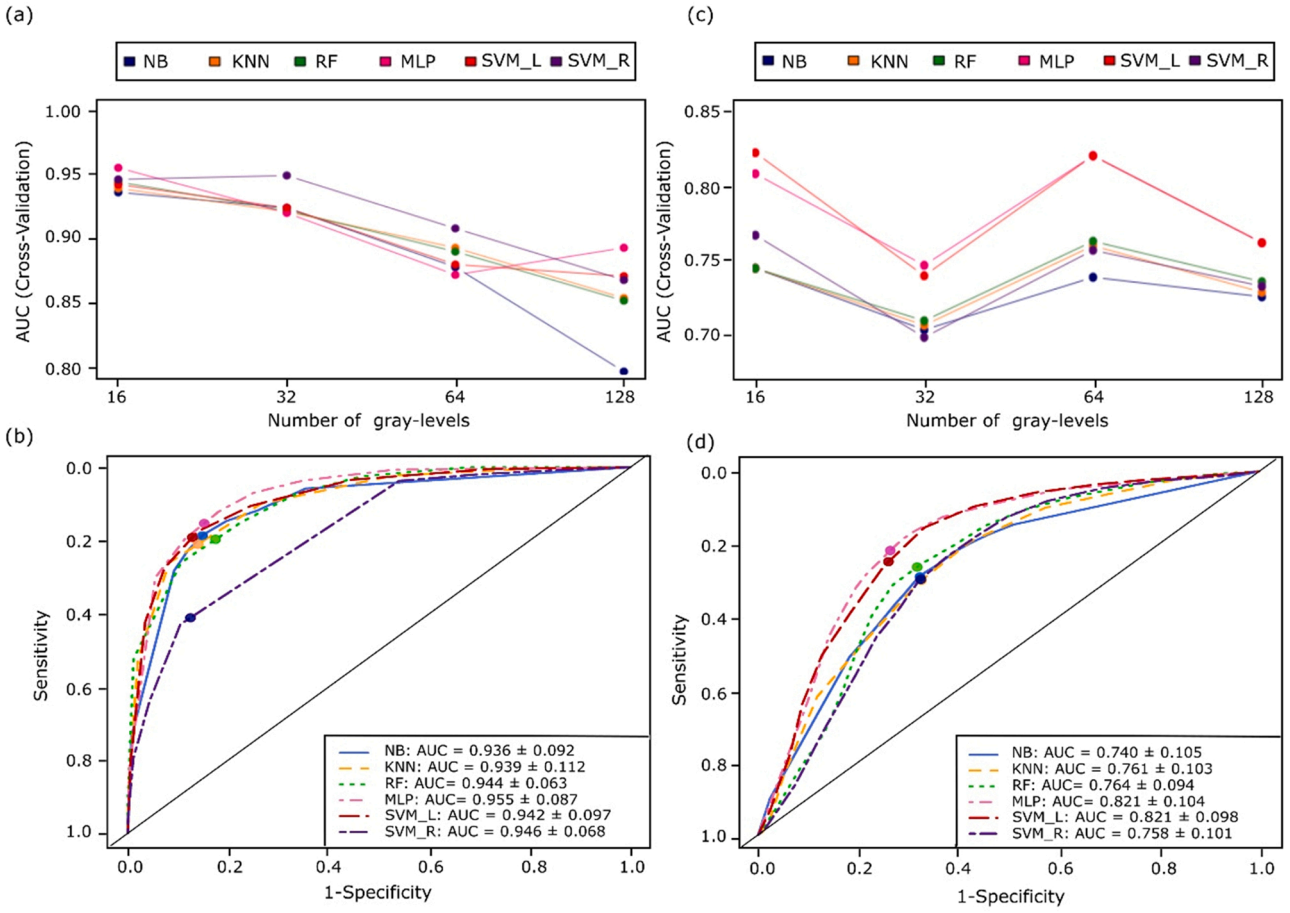

Fig. 7. Comparison among the maximum AUC results obtained using features extracted from the ROI quantized with four different nGL and applying the six models under study with the *p*-value feature selection method: (a) animal study, (c) human study. Average ROC curves for the six models using the *p*-value feature selection method: (b) animal study (nGL=16), (d) human study (nGL= 64). The highlighted points on the curves indicate the optimal cutoff points that weight sensitivity and specificity equally, measured with the 'closest-to-(0,1)' criterion.

features used to achieve the best performance was considered a key factor since reducing the number of features usually implies a less complex and more generalizable model. This happened in the human study using structural features, where the SVM model achieved the maximum AUC values again but using all the available features. However, similar outcomes were obtained with the MLP model using less than half of the features, thus becoming a better candidate to be generalizable.

This work is not the first attempt to identify alcohol-related brain changes using machine learning techniques (Cosa et al., 2017; Kamarajan et al., 2020). Previous works also applied these techniques to fMRI data to predict alcohol use severity or dependence (Fede et al., 2019; Zhu et al., 2018), but mainly using values of FC within and between resting-state networks. Previous studies, however, demonstrate an overall lack of consensus and/or reproducibility of reliable patterns of FC alterations across rs- fMRI studies (Badea et al., 2017; He et al., 2019; King et al., 2019). Therefore, a robust method based on standardized feature computation could represent a significant step forward for developing robust imaging biomarkers based on rs-fMRI. The outcome of the present study supports the hypothesis that standardized radiomic features (Zwanenburg et al., 2020) extracted from rs-fMRI can identify robust brain network alterations driven by alcohol drinking, an important step towards reproducible radiomics.

The presented methodology was validated using structural MRI, achieving a high classification performance in both animals (AUC $>$ 0.95) and humans (AUC $>$ 0.8). Again, the human findings mirrored those obtained in the animal study, achieving the best overall performance, in this case, using the MLP model in both species. This was the case, even considering that the structural contrast was different in both species, with T2 or T1 weighted images in rats and humans, respectively. This result highlights the robustness of radiomics for biomarker discovery. The use of combined features (functional and structural) was not evaluated as it was out of the scope of this work. As previously commented, one of our main objectives was to develop an alternative approach to identify potential AUD biomarkers using radiomic features instead of features related to FC, which are mainly used in these kinds of works. This process also allowed us to focus on identifying key brain networks in alcohol addiction and demonstrating the possibility of capturing functional heterogeneity. Since this imaging modality is commonly used in the study of AUD, our results now open the possibility of retrospective analysis of published data in which targeted ROIs can be selected based on functional (fMRI-based) findings. However, future work will evaluate the possibility of a combined analysis. Even so, it is worth mentioning that not always applying a multimodal analysis (i.e., combining functional and structural MRI) implies better results since the additional features may lead to overfitting in the training process (Fede et al., 2019).

We also proved that the discriminative power of the features was affected by the quantization process, obtaining different results depending on the nGL used. These results support previous studies that concluded that the optimal nGL should be determined according to the application (Mahmoud-Ghoneim et al., 2008; Ortiz-Ramón et al., 2018).

In this work, we compared two FS methods, a filter based on the *p*-value and a wrapper (SVM-RFE). In general, the best results were

obtained using the filter method. The only exception was in the animal study using functional features, where we considered the wrapper a better option because the model required fewer features, but there were no statistically significant differences between both results. We observed that the FS algorithm was relevant in the process as the best results were achieved with a substantially reduced set of features in all cases.

In all the subjects (animals and humans with both imaging modalities), most of the optimal features used by the models with better performance were derived from the GLSZM and the GLRLM feature extraction methods. These features provide regional information, defining the heterogeneity properties of the region by measuring the distribution of connected groups of pixels with identical gray-levels values. Therefore, the identified heterogeneity differences in the analyzed ROIs between AUD and controls depended mainly on regional gray-levels variations. Therefore, the results obtained show the ability of these features to capture functional heterogeneity, highlighting their possible role in capturing physiological events. Elucidating the nature of these events will require further work in animal models and/or post-mortem tissue to investigate the neurobiological underpinnings. Recently, a microglial reaction in the grey matter associated with a change in the extracellular space geometry (De Santis et al., 2020), as well as alterations in white matter microstructure (De Santis et al., 2019), were found in msP rats after one month of drinking and paralleled by DTI findings in AUD patients. These alterations might support the texture biomarkers identify in the current study.

Previous studies have demonstrated that features extracted from volumetric regions can capture more information about the region heterogeneity and are more discriminative than 2D features extracted from a single slice (Depeursinge et al., 2014; Ortiz-Ramón et al., 2018). Although structural MRI is the modality most used in radiomics analysis, extrapolating this approach to fMRI reveals an enormous potential in detecting AUD patients and identifying relevant brain networks in alcohol addiction. However, a lower performance using fMRI data was expected as parameters such as spatial resolution have been reported to have a particularly strong effect on radiomic features (Mayerhoefer et al., 2020). Discrimination based on texture analyses improves with higher spatial resolution since the finest textural details can be spotted, so we consider that the better resolution of the structural scans is an essential factor that could have an important impact on the difference in the results. Despite that, promising results have been achieved employing only functional data, and further validation would be necessary to ensure that the developed models are generalizable for both imaging modalities.

The fact that the findings in humans reflect those obtained in rats is important to establish a causal relationship between alcohol drinking and the detected brain network changes. Using an animal model allows us to avoid the comorbidities, the use of medication, and heterogeneous drinking patterns usually found in AUD patients, thus allowing causal inference on the effects of alcohol. However, we consider important to mention the different nature of alcohol effects on the brain by comparing exposures of 4 weeks (rats) and decades (humans). It may be the case that network effects that occur after four weeks of exposure change over years of exposure, ultimately resulting in structural changes. Indeed, these two conditions are not comparable when assessing tissue alterations, and our intention is not to draw this direct parallelism. However, the results in the animal model have the added value of supporting the utility of the developed technique to identify initial changes in the system.

Our study has several limitations. First, while in the rat study, the animals were evaluated before and immediately after 30 days of alcohol consumption, in the human study, healthy controls were compared with abstinent alcoholics (abstinence days: 21 $\pm$ 7). Therefore, more analyses are needed to clarify the contribution of abstinence to the found results. Second, the sample under study included only male participants, as most of the AUD patients admitted to our inpatient care are males. Therefore, future studies using gender-matched groups are necessary. Finally, although the presented methodology has been validated using another imaging modality, further validation using an independent cohort is required for future work.

5. Conclusion

The results presented in this study show that 3D radiomic features extracted from rs-fMRI combined with machine learning techniques can identify subjects with AUD in both rats and humans with good accuracy. The results obtained also confirm the striatal network as a key target of alcohol effects. This emphasizes the utility of the presented radiomics approach to identify relevant networks in drug addiction. Although further validation is required, the results obtained are promising and confirm the validity of radiomic features as potential AUD biomarkers.

CRediT authorship contribution statement

Silvia Ruiz-España: Writing – original draft, Methodology, Software, Formal analysis, Validation. **Rafael Ortiz-Ramón:** Software, Formal analysis. **Úrsula Pérez-Ramírez:** Software. **Antonio Diaz-Parra:** Software. **Roberto Ciccocioppo:** Resources. **Patrick Bach**: Resources. **Sabine Vollstädt-Klein:** Resources. **Falk Kiefer:** Resources. **Wolfgang H. Sommer:** Resources, Writing – review & editing, Funding acquisition. **Santiago Canals:** Writing – review & editing, Funding acquisition, Conceptualization, Resources, Supervision. **David Moratal:** Writing – review & editing, Funding acquisition, Conceptualization, Resources, Supervision.

Declaration of Competing Interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work was supported by the European Union's Horizon 2020 research and innovation program (668863-SyBil-AA) and the ERA-NET NEURON program (FKZ 01EW1112-TRANSALC and PIM2010ERN-00679), as well as the Spanish State Research Agency through the Severo Ochoa Program for Centres of Excellence in R&D (SEV- 2017–0723). S. C. acknowledges financial support from the Ministerio de Economía y Competitividad (MINECO) under grant PGC2018–101055-B-I00. D.M. and S.C. acknowledge financial support from the Generalitat Valenciana through the Prometeo Program (PROMETEO/2019/015). Additional support was given to W.H.S by the Deutsche Forschungsgemeinschaft Center grant TRR 265 (Heinz et al., 2020) and the Bundesministerium für Bildung und Forschung (BMBF; FKZ: 031L0190A, 01ZX1909CA). We thank Dr. Begoña Fernández for excellent technical assistance and Dr. Cécile Bordier for helping with the human data preprocessing.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.compmedimag.2023.102187.

References

Ambroise, C., McLachlan, G.J., 2002. Selection bias in gene extraction on the basis of microarray gene-expression data. Proc. Natl. Acad. Sci. U. S. A 99, 6562–6566. https://doi.org/10.1073/pnas.102102699.

Avanzo, M., et al., 2017. Beyond imaging: The promise of radiomics. Phys. Med. 38, 122–139. https://doi.org/10.1016/j.ejmp.2017.05.071.

Bach, P., et al., 2020. P.033 Attenuation of alcohol-cue modulated insula connectivity by naltrexone predicts relapse in alcohol-dependent patients: a longitudinal clinical neuroimaging trial. Eur. Neuropsychopharmacol. 40, S23. https://doi.org/10.1016/j.euroneuro.2020.09.036.
Bach, P., et al., 2021. FMRI-based prediction of naltrexone response in alcohol use disorder: a replication study. Eur. Arch. Psychiatry Clin. Neurosci. 271, 915–927. https://doi.org/10.1007/s00406-021-01259-7.
Badea, L., et al., 2017. Exploring the reproducibility of functional connectivity alterations in Parkinson's disease. PLoS One 12, e0188196. https://doi.org/10.1371/journal.pone.0188196.
Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging 23, 137–152. https://doi.org/10.1109/TMI.2003.822821.
Bordier, C., et al., 2021. Increased network centrality of the anterior insula in early abstinence from alcohol. Addict. Biol., e13096 https://doi.org/10.1111/ADB.13096.
Chanraud, S., et al., 2011. Disruption of functional connectivity of the default-mode network in alcoholism. Cereb. Cortex 21, 2272–2281. https://doi.org/10.1093/cercor/bhq297.
Ciccocioppo, R., et al., 2006. Genetically selected Marchigian Sardinian alcohol-preferring (msP) rats: An animal model to study the neurobiology of alcoholism. Addict. Biol. 11, 339–355. https://doi.org/10.1111/j.1369-1600.2006.00032.x.
Collewet, G., et al., 2004. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn. Reson. Imaging 22, 81–91. https://doi.org/10.1016/j.mri.2003.09.001.
Cosa, A., et al., 2017. Multi-modal MRI classifiers identify excessive alcohol consumption and treatment effects in the brain. Addict. Biol. 22, 1459–1472. https://doi.org/10.1111/adb.12418.
Crews, F., Nixon, K., 2009. Mechanisms of neurodegeneration and regeneration in alcoholism. Alcohol Alcohol 44, 115–127. https://doi.org/10.1093/ALCALC/AGN079.
Dadi, K., et al., 2019. Benchmarking functional connectome-based predictive models for resting-state fMRI. Neuroimage 192, 115–134. https://doi.org/10.1016/j.neuroimage.2019.02.062.
De Santis, S., et al., 2019. Microstructural White Matter Alterations in Men with Alcohol Use Disorder and Rats with Excessive Alcohol Consumption during Early Abstinence. JAMA Psychiatry 76, 749–758. https://doi.org/10.1001/jamapsychiatry.2019.0318.
De Santis, S., et al., 2020. Chronic alcohol consumption alters extracellular space geometry and transmitter diffusion in the brain. Sci. Adv. 6, eaba0154. https://doi.org/10.1126/sciadv.aba0154.
Depeursinge, A., et al., 2014. Three-dimensional solid texture analysis in biomedical imaging: Review and opportunities. Med Image Anal. 18, 176–196. https://doi.org/10.1016/j.media.2013.10.005.
Dupuy, M., Chanraud, S., 2016. Imaging the addicted brain: alcohol. Int Rev. Neurobiol. 129, 1–31. https://doi.org/10.1016/bs.irn.2016.04.003.
Fede, S.J., et al., 2019. Resting state connectivity best predicts alcohol use severity in moderate to heavy alcohol users. NeuroImage Clin. 22, 101782 https://doi.org/10.1016/j.nicl.2019.101782.
Feng, Q., Ding, Z., 2020. MRI Radiomics Classification and Prediction in Alzheimer's Disease and Mild Cognitive Impairment: A Review. Curr. Alzheimer Res. 17, 297–309. https://doi.org/10.2174/1567205017666200303105016.
Fernández-Delgado, M., et al., 2014. Do we need hundreds of classifiers to solve real world classification problems. J. Mach. Learn. Res 15, 3133–3181. https://doi.org/10.1016/j.csda.2008.10.033.
Gibbs, P., Turnbull, L.W., 2003. Textural analysis of contrast-enhanced MR images of the breast. Magn. Reson. Med. 50, 92–98. https://doi.org/10.1002/mrm.10496.
Gillies, R.J., et al., 2016. Radiomics: images are more than pictures, they are data. Radiology 278, 563–577. https://doi.org/10.1148/radiol.2015151169.
Guyon, I., et al., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46, 389–422. https://doi.org/Doi10.1023/A:1012487302797.
Hassan, I., et al., 2016. Radiomic Texture Analysis Mapping Predicts Areas of True Functional MRI Activity. Sci. Rep. 6, 25295. https://doi.org/10.1038/srep25295.
Jenkinson, M., et al., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841. https://doi.org/10.1016/S1053-8119(02)91132-8.
Jenkinson, M., et al., 2012. FSL. Neuroimage 62, 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015.
Juhás, M., et al., 2017. Deep grey matter iron accumulation in alcohol use disorder. Neuroimage 148, 115–122. https://doi.org/10.1016/J.NEUROIMAGE.2017.01.007.
Kalthoff, D., et al., 2011. Functional connectivity in the rat at 11.7T: Impact of physiological noise in resting state fMRI. Neuroimage 54, 2828–2839. https://doi.org/10.1016/j.neuroimage.2010.10.053.
Kamarajan, C., et al., 2020. Random forest classification of alcohol use disorder using EEG source functional connectivity, neuropsychological functioning, and impulsivity measures. Behav. Sci. (Basel). 10, 62. https://doi.org/10.3390/bs10030062.
Karahanoglu, F.I., Van De Ville, D., 2015. Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. Nat. Commun. 6, 1–10. https://doi.org/10.1038/ncomms8751.
King, J.B., et al., 2019. Generalizability and reproducibility of functional connectivity in autism. Mol. Autism 10, 27. https://doi.org/10.1186/s13229-019-0273-5.
Kohno, M., et al., 2017. Executive control and striatal resting-state network interact with risk factors to influence treatment outcomes in alcohol-use disorder. Front. Psychiatry 8, 182. https://doi.org/10.3389/fpsyt.2017.00182.
Kuhn, M., 2008. Building predictive models in R using caret package. J. Stat. Softw. 28, 1–26. https://doi.org/10.1053/j.sodo.2009.03.002.
Kuhn, M., Johnson, K., 2013. An introduction to feature selection. Appl. Predict. Model. 487–519. https://doi.org/10.1007/978-1-4614-6849-3_19.
Liu, T.T., et al., 2017. The Global Signal in fMRI: Nuisance or Information. Neuroimage 150, 213–229. https://doi.org/10.1016/j.neuroimage.2017.02.036.
Liu, Z., et al., 2019. The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. Theranostics 9, 1303–1322. https://doi.org/10.7150/thno.30309.
Mahmoud-Ghoneim, D., et al., 2008. The impact of image dynamic range on texture classification of brain white matter. BMC Med. Imaging 8, 18. https://doi.org/10.1186/1471-2342-8-18.
Mayerhoefer, M.E., et al., 2020. Introduction to Radiomics. J. Nucl. Med 61, 488–495. https://doi.org/10.2967/JNUMED.118.222893.
Mikl, M., et al., 2008. Effects of spatial smoothing on fMRI group inferences. Magn. Reson. Imaging 26, 490–503. https://doi.org/10.1016/j.mri.2007.08.006.
Müller-Oehring, E.M., et al., 2015. The resting brain of alcoholics. Cereb. Cortex 25, 4155–4168. https://doi.org/10.1093/cercor/bhu134.
Ortiz-Ramón, R., et al., 2018. Classifying brain metastases by their primary site of origin using a radiomics approach based on texture analysis: a feasibility study. Eur. Radiol. 28, 4514–4523. https://doi.org/10.1007/s00330-018-5463-6.
Pan, W.J., et al., 2013. Infraslow LFP correlates to resting-state fMRI BOLD signals. Neuroimage 74, 288–297. https://doi.org/10.1016/j.neuroimage.2013.02.035.
Pan, W.J., et al., 2015. Considerations for resting state functional MRI and functional connectivity studies in rodents. ront Neurosci. 9, 269. https://doi.org/10.3389/fnins.2015.00269.
Park, Y.W., et al., 2020. Differentiating patients with schizophrenia from healthy controls by hippocampal subfields using radiomics. Schizophr. Res. 223, 337–344. https://doi.org/10.1016/j.schres.2020.09.009.
Perez-Ramirez, U., et al., 2017. Brain functional connectivity alterations in a rat model of excessive alcohol drinking: A resting-state network analysis. Conf. Proc. IEEE Eng. Med Biol. Soc. (EMBC) 2017, 3016–3019. https://doi.org/10.1109/EMBC.2017.8037492.
Pfefferbaum, A., et al., 2010. Diffusion tensor imaging of deep gray matter brain structures: Effects of age and iron concentration. Neurobiol. Aging 31, 482–493. https://doi.org/10.1016/J.NEUROBIOLAGING.2008.04.013.
Rizzo, S., et al., 2018. Radiomics: the facts and the challenges of image analysis. Eur. Radiol. 2, 36. https://doi.org/10.1186/s41747-018-0068-z.
Rouach, H., et al., 1990. Effect of acute ethanol administration on the subcellular distribution of iron in rat liver and cerebellum. Biochem. Pharmacol. 39, 1095–1100. https://doi.org/10.1016/0006-2952(90)90289-W.
Schwarz, A.J., et al., 2006. A stereotaxic MRI template set for the rat brain with tissue class distribution maps and co-registered anatomical atlas: Application to pharmacological MRI. Neuroimage 32, 538–550. https://doi.org/10.1016/j.neuroimage.2006.04.214.
Scuppa, G., et al., 2020. Aberrant insular cortex connectivity in abstinent alcohol-dependent rats is reversed by dopamine D3 receptor blockade. Addict. Biol. 25, e12744 https://doi.org/10.1111/ADB.12744.
Shu, Z., et al., 2021. Predicting the progression of Parkinson's disease using conventional MRI and machine learning: An application of radiomic biomarkers in whole-brain white matter. Magn. Reson. Med. 85, 1611–1624. https://doi.org/10.1002/mrm.28522.
Sladky, R., et al., 2011. Slice-timing effects and their correction in functional MRI. Neuroimage 58, 588–594. https://doi.org/10.1016/j.neuroimage.2011.06.078.
Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155. https://doi.org/10.1002/hbm.10062.
Smith, S.M., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23, S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051.
van Timmeren, J.E., et al., 2020. Radiomics in medical imaging—"how-to" guide and critical reflection. Insights Imaging 11, 91. https://doi.org/10.1186/s13244-020-00887-2.
Vallières, M., et al., 2015. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys. Med. Biol. 60, 5471–5496. https://doi.org/10.1088/0031-9155/60/14/5471.
Van Buuren, M., et al., 2009. Cardiorespiratory effects on default-mode network activity as measured with fMRI. Hum. Brain Mapp. 30, 3031–3042. https://doi.org/10.1002/hbm.20729.
Waugh, S.A., et al., 2011. The influence of field strength and different clinical breast MRI protocols on the outcome of texture analysis using foam phantoms. Med. Phys. 38, 5058–5066. https://doi.org/10.1118/1.3622605.
Weiland, B.J., et al., 2014. Reduced left executive control network functional connectivity is associated with alcohol use disorders. Alcohol. Clin. Exp. Res. 38, 2445–2453. https://doi.org/10.1111/acer.12505.
World Health Organization, 2018, Global status report on alcohol and health 2018 (Geneva). ⟨https://www.who.int/publications/i/item/9789241565639⟩ (accessed 12.15.20).
Zhu, X., et al., 2017. Model-free functional connectivity and impulsivity correlates of alcohol dependence: a resting-state study. Addict. Biol. 22, 206–217. https://doi.org/10.1111/adb.12272.
Zhu, X., et al., 2018. Random forest based classification of alcohol dependence patients and healthy controls using resting state MRI. Neurosci. Lett. 676, 27–33. https://doi.org/10.1016/j.neulet.2018.04.007.
Zwanenburg, A., et al., 2020. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 295, 328–338. https://doi.org/10.1148/radiol.2020191145.