Psychiatry Research: Neuroimaging 362 (2026) 112255

Contents lists available at ScienceDirect

Psychiatry Research: Neuroimaging

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

Brain Connectivity and Machine Learning Approaches to assess the underlying neurobiology and prediction accuracy of anorexia nervosa: A replication study

Laura Monteiro Rente Dias [a,b], Hugo Schnack [a,c], Daniel Geisler [d], Marcel Reinders [b], Tonya White [a,e,1,*], Gwen Dieleman [a,1], Xucong Zhang [b,1]

[a] Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC-Sophia Children's Hospital, Rotterdam, the Netherlands
[b] DelftUniversity of Technology, the Netherlands
[c] Institute of Language Sciences, Utrecht University, Utrecht, the Netherlands
[d] Translational Developmental Neuroscience Section, Department of Child and Adolescent Psychiatry, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany
[e] Section on Social and Cognitive Developmental Neuroscience, National Institute of Mental Health, Bethesda, MD, USA

ARTICLE INFO

Keywords:
Anorexia nervosa
Graph theory
Machine Learning
Functional connectivity

ABSTRACT

Resting-state fMRI has been used to study aberrant functional connectivity properties in patients with anorexia nervosa (AN) at several stages of the illness. One popular way to extract these metrics is to use graph theory to showcase aberrant brain connectivity between patients with AN versus controls. However, most studies use classic analyses to investigate these differences, which could limit the number and choices of features used in one model. Instead, machine learning models have proven to be a promising tool in studying the functional connectivity of various disorders. In this study, we employ a combination of local graph metrics and a support vector machine to distinguish between first-onset AN (N = 56) cases and controls (N = 64). We replicate and extend prior work evaluating the predictive value of an existing machine learning approaches in detecting functional connectivity differences in patients with AN. Our method achieves an average classification accuracy of 65% with cross-validation evaluation. We further demonstrate that the results are driven mainly by the participation index of the nodes that are implicated in distinguishing the two groups. Our findings contribute to the growing body of evidence supporting the predictive value of resting-state fMRI in the study of anorexia nervosa.

Introduction

Anorexia nervosa (AN) is a severe psychiatric disorder with a peak onset in adolescence that pri-marily affects girls and young women, and carries one of the highest morbidity and mortality rates of all mental illnesses (Papadopoulos et al. 2009; Schmidt et al. 2016; Van Eeden et al., 2021). Previous studies have pointed to a complex heterogeneous etiology involving a combi-nation of neurobiological and psychosocial factors (Kaye and Christina E, 2013; Kaye et al., 2009). Despite this heterogeneity in profile, there seem to be underlying key features that characterize the population, such as intense fear of weight gain, restriction of caloric intake, and distorted body image (Neale and Hudson, 2020).

The pursuit of brain biomarkers has become an important area of interest in the fields of psychi-atry, psychology, and neurodevelopment. A substantial body of research has sought to accurately classify patients with a variety of disorders, using neuroimaging techniques and machine learning, that is based on the assumption that the distinctive symptoms observed in these populations have a neurobiological basis and that differentiation of these groups on the basis of such biomarkers is feasible (Arbabshirani et al. 2017). This approach has been successful for the classification of disorders such as depression, autism spectrum disorder, anxiety, and schizophrenia with some success (Khosla et al. 2019; Wen et al. 2024). The majority of research on AN has focused on utilizing traditional statistical techniques and neuroimaging to identify brain biomarkers. While there are significant statistical findings at the

* Corresponding author.
E-mail address: tonya.white@nih.gov (T. White).
[1] Authors contributed equally.

https://doi.org/10.1016/j.pscychresns.2026.112255
Received 18 July 2025; Received in revised form 18 May 2026; Accepted 20 May 2026
Available online 1 June 2026

group level in these studies, machine learning classifiers could facilitate enhanced sensitivity by aggregating more information and considering smaller effect sizes across a broader range of voxels, as well as allowing for multivariate analysis, which could be particularly interesting in AN due to its clinical and symptomatic heterogeneity Pereira et al., 2009).

A number of studies have employed both structural and functional neuroimaging techniques to gain insight into the neurobiological underpinnings of AN. Structural neuroimaging studies have predominantly identified significant reductions in global gray matter and cortical thinning. These alterations are primarily identified in acutely underweight patients and tend to normalize following successful weight recovery (King et al. 2018 Walton et al. 2022). Therefore, structural differences may be more closely associated with undernutrition than with the core behavioral symptoms of AN. Functional magnetic resonance imaging (fMRI) has been employed to assess alterations in brain connectivity, which may be independent of weight, and associated with key features of AN (Geisler et al. 2016). In fMRI, distinct brain activation patterns are captured by measuring the variance in blood-oxygen-level-dependent (BOLD) signal resulting from varying levels of brain activity (Lee et al., 2013). This metric can be used to obtain a measure of functional connectivity (FC) between different brain regions with FC being defined as the temporal dependence of neuronal activation patterns of distinct brain regions (White and Muetzel, 2014). By examining the interactions between brain regions, it is possible to determine characteristics of the brain at both regional and global levels (Stevens 2016).

Resting-state fMRI (rs-fMRI) has gained considerable popularity as a means of obtaining re-producible results pertaining to brain connectivity profiles as applied to the study of a number of disorders (Lee et al., 2013). In general, studies using rs-fMRI have typically applied either an independent component analysis (ICA), or seed-based connectivity analyses to identify differences in established brain networks, including the Default Mode Network (DMN) and the Executive Control Frontoparietal Network (FPN).

In the case of the study of FC alterations in AN, studies using rs-fMRI and the methods mentioned above have reported differences within functional areas related to reward and corticolimbic regulation (Gaudio, Wiemerslage, et al. 2016). A study by Boehm et al. (2014) reported increased connectivity between the frontoparietal network and the angular gyrus in adolescent females with AN. Furthermore, increased FC was observed between the DMN and the left anterior insula, a brain region linked to interoceptive awareness, which may be related to ruminations about bodily appearance and food (Boehm et al., 2014). Gaudio and Piervincenzi (2015) found decreased FC between the anterior cingulate cortex and the executive control network in female adolescents who are in the initial stages of AN. Moreover, Scaife et al. (2017) examined the neural correlates of somatosensory and interoceptive processes in individuals with AN and those who have recovered from the disorder. The application of whole-brain ICA revealed a decrease in FC across multiple networks, including the visual and auditory network, in both individuals with a current diagnosis of AN as well as recovered individuals (Scaife et al., 2017). Interestingly, in both recovered and AN patients, decreased temporal coherence within the lateral visual and auditory networks was found, specifically in the insula, compared to typically developed (TD) participants. Haynos et al. (2018) found a decrease in FC between the nucleus accumbens, and the superior frontal gyrus, between the ventral caudate and frontal and posterior regions, and between the dorsal caudate and frontal, temporal, and posterior regions in the AN group compared to TD.

In addition to ICA and region of interest approaches, studies have used graph theory to quantify alterations in functional connectivity at both global and local levels. The application of graph metrics enable a more nuanced examination of brain connectivity, moving beyond a mere focus on the synchronous activity of specific regions or components of interest. This approach allows for a more complex understanding of the topology and structure of brain connectivity (Geisler et al. 2020). Gaudio and Olivo (2018) used both global and local graph metrics to investigate differences in resting-state FC in a sample of adolescents at the earliest stage of AN, defined as less than six months. The study revealed no differences in either local or global metrics between the AN and TD groups when a seven-node network was employed, comprising the left and right rostral ACC, the left medial fronto-orbital gyrus, the left paracentral lobule, the left 10th lobule of the cerebellum, the right superior occipital gyrus, and the left posterior insula, to create weighted, undirected graphs (Gaudio and Olivo, 2018).

In another graph theory study, Geisler et al. 2016 employed seven global graph metrics to examine alterations in network configuration in individuals with acute AN. The findings revealed a reduction in connectivity strength and an increased path length, meaning the average route between nodes, in the posterior insula and thalamus in individuals with AN. The diminished local network efficiency observed in the thalamus and posterior insula may be attributed to impaired integration of visuospatial and homeostatic signals in this population (Gaudio et al. 2016; Geisler et al. 2016).

Interestingly, studies focusing on individuals who have recovered from AN have also reported alterations that appear to persist after weight recovery. These findings suggest the potential of a trait marker of the disorder that is independent of weight. Cowdrey et al. found increased connectivity between the dorsolateral prefrontal cortex/inferior frontal gyrus and precuneus in individuals who had recovered from AN (Cowdrey et al., 2014). Similarly, Boehm et al. 2016 reported alterations in the frontoparietal network, which has been associated with cognitive control. The result of the study by Geisler et al. 2020 indicated a globally disrupted brain network architecture, which could be a sign of an alteration present prior to undernutrition that does not normalize with recovery. As an exploratory analysis, the authors applied a support vector machine (SVM), a machine learning algorithm for classification problems, to the local graph metrics. This resulted in a classification score of 70.4% with a p value of 0.0003, in the task of classifying recovered AN participants from TD, thereby demonstrating the potential for predictive classification of patients at an individual level using brain-based biomarkers (Geisler et al. 2020). However, no assessment of the generalizability of the model on an outer sample was performed.

In order to replicate the findings of Geisler et al. 2020, we perused the development of an SVM method for the classification of AN patients employing graph metrics derived from rs-fMRI connectivity, within the BRAVE study, a large first-onset AN longitudinal case-control study in the Netherlands, performed at the Erasmus Medical Centre (Erasmus MC). Our primary goal was to use an independent sample of AN patients to replicate the study of Geisler et al. Geisler et al. 2020. We set out to assess whether using their method in an independent sample, it would be feasible to develop a model for the classification of individuals with AN based on their functional connectivity. We hypothesize that the key features of AN could be related to a characteristic brain pattern or biomarker, present in all patients with AN, despite their heterogeneous symptoms and profile.

Methods

Our method draws on the approach outlined in Geisler et al. 2020 which rep-resents one of the few instances where machine learning has been employed for the classification of AN using fMRI data. An overview of our method is shown in (1). Briefly, from preprocessed fMRI data, we extracted time series and calculated connectivity matrices from 160 spherical nodes for each participant to form a graph. The local graph metrics are then extracted and fed into a linear SVM to perform classification between two groups: one with first-onset AN participants, and the other comprising of TD participants. We further analyze the nodes that affects classification accuracy in the SVM to demonstrate the consistency between our findings and previous works.

BRAVE study design

The BRAVE study is a longitudinal first-onset AN case-control study with the objective of identify-ing an association between clinically significant changes in symptoms and the underlying behavioral, neurobiological, cognitive, and physical health changes. The data used in this work corresponds to the initial acquisition point for both the typically developed and anorexia nervosa groups (Steegers et al. 2024). The BRAVE study was approved by the Medical Ethical Committee of the Erasmus Med-ical Centre – Sophia Children's hospital (Erasmus MC-Sophia) (MEC 2016–194/NL55175.078.16) and was conducted in agreement with the Declaration of Helsinki.

Participants

The AN group comprised 56 female participants with a first-onset typical or atypical AN diagnosis according to the DSM 5 classification shown in Table 1 (American Psychiatric Association, DSMTF, DS American Psychiatric Association 2013), and the TD group consisted of 64 typically developing female participants whose age and education level were matched to the AN group. For the AN group, participants were included if they were female, aged between 12 and 22 years old, and diagnosed with first-onset typical AN or atypical AN according to DSM-5 criteria within 12 months prior to inclusion. For the TD group, participants were required to have a healthy body weight, defined as a body mass index – standard deviation score (BMI-SDS) between −1.3 and +1.3. The exclusion criteria for both groups included the presence of a psychotic, neurological, or substance abuse disorder, severe motor and sensory impairments, IQ <70 as measured by an intelligence test, and insufficient Dutch language skills (Steegers et al. 2024).

Table 1
DSM-5-TR Criteria for Typical and Atypical Anorexia Nervosa (American Psychiatric Association, DSMTF, DS American Psychiatric Association 2013).

DSM-5-TR Criterion	Typical Anorexia Nervosa	Atypical Anorexia Nervosa
Criterion A: Restriction of energy intake Body Weight Status		
Criterion B: Intense fear of weight gain		
Criterion C: Disturbance in self-perceived weight or shape		
Persistent restriction of energy intake leading to significantly low body weight relative to age, sex, developmental trajectory, and physical health.		
Significantly low body weight (e.g., BMI below minimally normal level for adults; below expected for children/adolescents).		
Intense fear of gaining weight or becoming fat, or persistent behavior that interferes with weight gain, even though underweight.		
Disturbance in the way body weight/shape is experienced; undue influence of weight/shape on self-evaluation; or persistent lack of recognition of seriousness of low body weight.		
All criteria for AN are met except weight remains within or above the normal range despite significant weight loss.		
Weight is not significantly low, but the individual may have experienced substantial weight loss.		
Same as typical AN — intense fear of weight gain or persistent behaviors preventing weight gain.		
Same cognitive disturbance present despite normal or above-normal weight.		
Subtypes eating/Purging type specified. May also present with restricting or binge-eating/purging behaviors.	Restricting type or Binge-	

MRI data acquisition

MRI acquisitions were performed using a 3.0 tesla system (GE Discovery 750 w, GE Healthcare, Milwaukee, WI, USA) with an 8-channel head coil. Structural T1-weighted images were acquired using an inversion recovery fast spoiled gradient recalled (IR-FSPGR) sequence, with the following parameters: TR=8.77 ms, TE=3.4 ms, TI=600 ms, flip angle=10°, acquisition matrix=220*pixels* × 220*pixels*, field of view (FOV)=220 *mm* × 220 *mm*, and slice thickness=1.0 mm. Echo planar imaging was employed for the rs-fMRI session with the following parameters: TR=1760 ms, TE=30 ms, flip angle=85°, acquisition matrix=64 × 64 pixels, FOV=230 *mm* × 230 *mm*, slice thickness=4 mm, volume=200. Subsequent to the MRI acquisition, all scans were evaluated for both data quality and incidental findings.

MRI data preprocessing and data quality

All rs-fMRI images were preprocessed using FSL's FMRIB's Software Library and Python 3.6. First, we cleaned each dataset by removing potential biases resulting from subject motion, car-diac/respiratory physiology, and scanner noise using MELODIC ICA data exploration with FSL's fMRI Expert Analysis Tool (FEAT). The first four volumes of each subject dataset were removed, trimming the scans to 196 vol each. Next, the single-session independent components were manually classified as either signal or noise. Following this procedure, components of non-interest were removed from each participant file by regressing out those time series as a first step to ensure quality.

After denoising, slice timing correction, motion correction, brain extraction, and image regis-tration were performed. For motion correction, mcFLIRT was used, and motion parameter ma-trices were obtained for further use. Image registration was performed to fit the data into an age-appropriate, T1-weighted standard space using a two-stage process. First, the fMRI data was registered to each participant's T1-weighted anatomical image. Next, the anatomical T1-weighted image was registered to the age-appropriate 2 mm standard MNI space T1-weighted template. Fi-nally, the resulting matrix from the concatenation of the previous two steps was applied to the denoised functional image. These steps were performed and concatenated using FSL's FEAT with 12 DOF.

The output matrices from the motion correction step with the translation coordinates were used to identify participants whose data was corrupted by motion. Participants were excluded if the mean root-mean-squared relative translation motion was greater than 0.2 mm and if the root-mean-squared maximum translation exceeded 1 mm. In the end, only one AN participant was excluded due to excessive motion (root-mean-squared maximum = 3.7,mean root-mean-squared relative translation
= 0.06).

Computation of graph metrics

In order to perform the classification between AN and TD participants based on functional connec-tivity, local functional graph metrics were computed and used as features for the model. Firstly, 160 spherical nodes with a diameter of 10 mm were centered at the voxel coordinates outlined in Geisler's work (Geisler et al. 2020). For each participant, average time series were extracted from within the boundaries of each sphere and Pearson correlation matrices between all pairs of nodes were calculated.

Based on the absolute values of the correlation matrices, weighted undirected graphs with 160 nodes were constructed for each subject using the open-source Python library Networkx, version

3.2.1 for Python 3.9[1]. The graphs were then thresholded with 21 sparsity levels removing between 70 - 90% of the edges (Geisler et al. 2016). Local graph metrics that characterize each node individually were calculated for each sparsity level for each participant.

Local metrics, meaning node metrics, include (1) degree, which corresponds to the number of edges adjacent to the node; (2) strength, which is the sum of weights of edges from neighboring nodes connected to the node; (3) local characteristic pathlength (CPLloc), which is the number edges in the shortest paths to any other node in the graph, measuring the integration of information; (4) betweenness centrality index (BCI), which is the fraction of all shortest paths in the network that contain the node of interest; (5) participation index (PI), which reflects how evenly distributed a node's edges are in a given community; (6) local efficiency (Eloc), which describes the efficiency of information transfer between the node of interest and all neighboring nodes by calculating the global efficiency of each neighborhood, meaning the multiplicative inverse of the shortest path distance between the nodes; (7) local clustering coefficient (CCloc) and (8) local efficiency normalized by the global efficiency of the whole graph (LEGE) (Geisler et al. 2020). If a metric is related to a path, such as CPLloc and BCI, the weight of the graph was not defined in terms of the correlation between two nodes, but rather as a distance, which corresponds to the inverse of the correlation value.

Subsequently, the area under the curve (AUC) across all 21 thresholds was calculated using a trapezoidal algorithm. This was done in order to obtain local graph features that were independent of the threshold used to sparsify the graphs. The AUC of each of the eight metrics at each of the 160 nodes was used as the input feature set for our model.

Support vector machine

After the extraction of the features of interest, namely the AUC of all local graph metrics, they were normalized across all subjects to comprise values between 0 and 1. After normalization, the method applied by Geisler et al. 2020 was replicated, which uses a linear SVM to classify between first on-set AN and TD participants (Model 1). The SVM is a supervised learning algorithm that computes hyperplanes in a multidimensional space that separate different classes (Ghosh 2019). An L2 regularization was applied with a regularization strength parameter set to C = 1. For evaluation,a k-fold cross-validation strategy was used with k

= 10. This means that the model is trained on $k - 1$ of the folds and validated on the remaining one fold. To assess the significance of the performance of the classifier, permutation testing was employed with 3000 permutations. This process involves repeating the classification with random iterations of the label vector, thereby testing the null hypothesis that the obtained classification was obtained by chance Fig. 1.

It should be noted that the aforementioned method did not include hyperparameter tuning, and that the value of C used may not be optimal for the data in question. Therefore, a linear SVM approach was repeated, while a nested cross-validation method was applied, using a grid search on the outer layer for hyperparameter tuning with C=$\{1e-8, 1.6-7, 2.5e-6, 4e-5, 6.3e-4, 1e-2, 1.6e-1, 2.5, 4e1, 6.3e2, 1e4\}$. The number of folds ($k$) were set to 5 for both the outer and inner cross-validations, in order to identify optimal values of C (Model 2). A p value was obtained through permutation testing to evaluate the validity of the classification.

In order to assess the impact of the discrepancies in population characteristics between our data set and that of Geisler et al. Geisler et al. 2020 on the classification, we trained our SVM model using only those features that exhibited the greatest positive and negative weights in Geisler et al. 2020 results (Model 3).

[1]https://networkx.github.io.

Results

Sample characteristics

Table 2 shows the sample characteristics of the AN and TD groups in the BRAVE study that were included Steegers et al. 2024. No differences were observed between the two groups with respect to age at MRI acquisition, ethnicity, maternal education, IQ, and medication. However, as expected, the AN group had a lower BMI-SDS and more eating disorder, depressive, and anxiety symptoms, as well as a higher prevalence of anxiety and mood disorders compared to the TD group. For further details regarding the BRAVE sample, please refer to Steegers et al. 2024.

Classification

After we calculated the 160 node correlation graph metrics from the rs-fMRI images, in the BRAVE study, we trained a classifier to differentiate between AN and TD participants, which is evaluated in a cross-validated way as well as compared to a classifier with random labels. In order to provide a point of comparison, the performance of the original model by Geisler et al. 2020, which did not include nested cross-validation and hyperparameter tuning, was also evaluated in the BRAVE sample. This resulted in a classification of 0.633 with a p value of 0.01 as shown in Fig. 2.

After using grid search nested cross-validation to find the

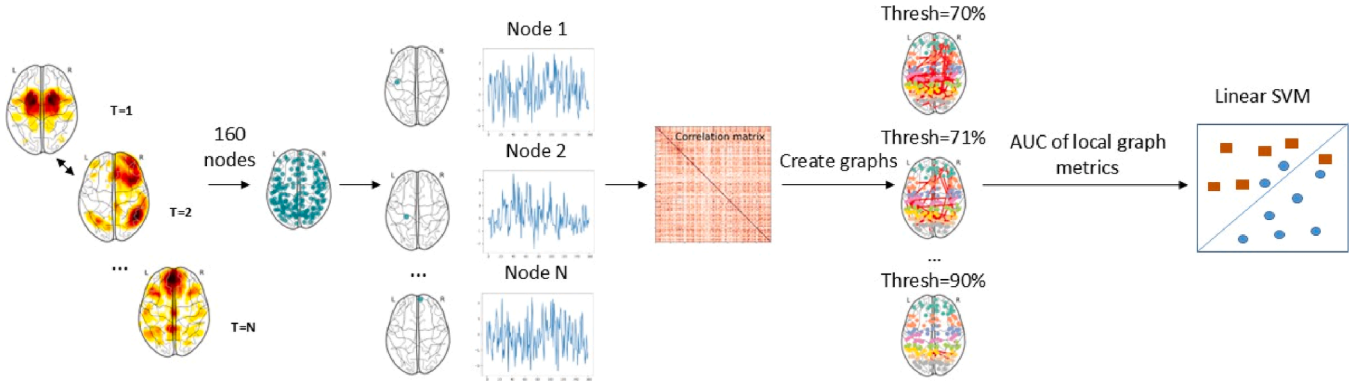
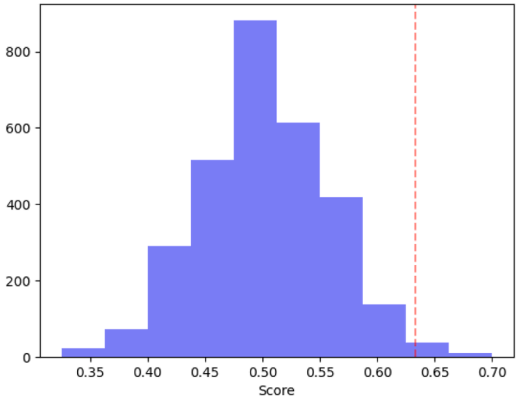


Fig. 1. Flowchart illustrating the methodology used to extract features from the fMRI image and perform classification using local graph metrics. From the processed fMRI data, we extracted the time series of 160 spherical nodes. With these, we calculated connectivity matrices for each participant, and subsequently created and thresholded graphs. We then extracted local graph metrics and used them as features for a linear SVM classifier.

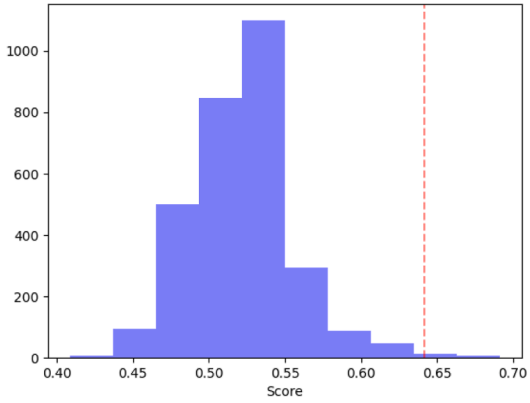
Table 2
Sociodemographic and clinical characteristics of the AN and TD participants.

	N	Statistic	AN participants (n = 56)	TD participants (n = 64)	p value
Age at MRI acquisition (years)	120	Median (IQR)	16.3 (15.5, 18.4)	17.3 (16.1, 18.3)	.06
Maximum movement	120	Mean (SD)	0.14 (0.05)	0.17 (0.07)	.01*
Mean movement	120	Mean (SD)	0.013 (0.009)	0.014 (0.013)	.7
Ethnicity	117	Percentage			.88
Dutch			96.3	96.8	
Western			1.9	1.6	
Non Western			1.9	1.6	
SES	113	Percentage			.14
Low			13.0	6.8	
Middle			37.0	30.5	
High			50.0	62.7	
IQ score	120	Mean (SD)	109.8 (12.9)	111.8 (12.0)	.39
BMI-SDS	119	Mean (SD)	−1.3 (1.3)	0.5 (1.0)	<*0*.001***
BMI-SDS <1.3	30	Percentage	53.6	-	
BMI-SDS >1.3	26	Percentage	46.4	-	
Psychotropic medication (% using)	120	Percentage			
Antidepressant			5.4	1.6	.25
Anxiolytic			3.6	1.6	.48
Antipsychotic			12.5	0	.004**
Stimulant			0	6.3	.06
EDE total score	120	Median (IQR)	3.8 (2.6, 4.6)	0.2 (0.1, 0.5)	<*0*.001***
Duration of illness (months)	49	Median (IQR)	3.8 (1.7, 7.4)	-	-
BDI total score	109	Median (IQR)	29.0 (20.0, 42.5)	4.0 (1.0, 7.75)	<*0*.001***
Any mood disorder	118	Percentage	58.2	14.3	<*0*.001***
SCARED total score	110	Median (IQR)	44.5 (28.0, 68.5)	21.5 (13.25, 32.25)	<*0*.001***
Any anxiety disorder	118	Percentage	45.5	12.7	<*0*.001***

Note - AN: anorexia nervosa; BDI-II: Beck Depression Inventory – Second edition; BMI-SDS: body mass index – standard deviation score; EDE: Eating Disorder Examination; IQ: Intelligence Quotient; MRI: Magnetic Resonance Imaging; SES: Socioeconomic status based on maternal education; categorized into: Low: primary education and lower general secondary education; Middle: higher general secondary education; High: higher vocational secondary education and higher academic education; SCARED: Screening for Child Anxiety Related Emotional Disorders; TD: typically developing. *p < .05, ** p < 0.01, *** p < .001.



(a) Results of the linear SVM without nested cross-validation and using all participants and all features. The histogram shows the classification using permutation (blue bars), and the actual classification as a red line.



(b) Results of the linear SVM with nested cross validation and using all participants and all features. The histogram shows the classification using permutation (blue bars), and the actual classification as a red line.

Fig. 2. Results of all SVM machine classifications. Results of the linear SVM without nested cross-validation and using all participants and all features. The histogram shows the classifi Results of the linear SVM with nested cross validation and using all participants and all fea-tures. The histogram shows the classification us-ing permutation (blue bars), and the actual clas-sification as a red line.

hyperparameters of the same linear model, we obtained a classification accuracy of 0.642 ± 0.057, over the 5 outer folds, with a *p* value of 0.007, with random permutations. Table 3 shows the average accuracy, F1 score, and specificity of the tuned linear model. Fig. 2b depicts the permutation results and the significance of the averaged cross-validation after applying the base method directly and with minor alterations. The 20 features with the most positive and negative weights of this classification are presented in Table 4 and illustrated in Fig. 3, which depicts their localization on the brain.

As an exploratory analysis, the linear classifier was run with optimal parameters but only with the most positive and negative weighted features from Geisler et al. 2020. This combination achieved a classification accuracy of 0.45 ± 0.103.

Table 3
Results on the linear SVM classification with nested cross-validation, including specificity, sensitivity, and f1-score for the classification of AN and TD groups. Values are reported as mean ± standard deviation across cross-validation folds.

Average accuracy	Average F1 score	Average specificity	Average sensitivity	
Model 1	0.633 ± 0.125	0.624 ± 0.135	0.714 ± 0.141	0.597 ± 0.252
Model 2	0.642 ± 0.057	0.626 ± 0.031	0.672 ± 0.043	0.641 ± 0.065
Model 3	0.450 ± 0.103	0.312 ± 0.201	0.543 ± 0.099	0.306 ± 0.214

Table 4
Features of support vector classifier (SVC): Only the features with the most positive and negative weights are presented. The name of the feature consists of the name of the brain region of the node, which is identified by a unique number as in Geisler, Borchardt, Boehm, et al. 2020 with the corresponding graph local feature. For more information regarding each node, please refer to the Supplementary of Geisler, Borchardt, Boehm, et al. 2020.

10 most positive features	SVM coefficients 1-midinsula L 56 LEGE		0.1170
2-	parietal L 54°	0.1174	
3-	postoccipital R 159 CCloc	0.1199	
4-ACC L 19 PI	0.1219		
5-	postinsula L 76 PI	0.1248	
6-	occipital L 141 PI	0.1291	
7-	temporal R 60 PI	0.1325	
8-	parietal L 54 PI	0.1578	
9-	vPFC L 23 PI	0.1620	
10-	postcingulate L 90 PI	0.1633	
10 most negative features			
11-	inftemporal L 91 PI	−0.1982	
12-dlPFC R 22 PI	−0.1865		
13-	supfrontal L 20 PI	−0.1665	
14-	parietal R 65 PI	−0.1600	
15-	latcerebellum L 98°	−0.1599	
16-	postcingulate L 80°	−0.1527	
17-	precuneus R 132 PI	−0.1506	
18-	antinsula L 28 PI	−0.1291	
19-	precentralgyrus R 51 CCloc	−0.1274	
20-vFC L 37 PI	−0.1270		

Analysis of features

Fig. 3.

Discussion

The main objective of this study was to evaluate the predictive capacity of an existing method, using graph metrics derived from rsfMRI images to differentiate first-onset AN patients from TD participants. While the methodology had previously been employed in a sample of recovered females with AN (Geisler et al. 2020), our study focused on a sample of first-onset AN in females aged 12 to 22 years old to ascertain the method's generalizability. We have a larger sample size when compared to other studies focused on AN, as well as different characteristics of our population, such as the disease stage of the participants and the matched TD group. This makes the approach a unique application of machine learning in the study of AN.

Our findings demonstrate that the implemented model effectively classified samples into AN or TD based on local graph metrics, achieving an accuracy of 65% which is slightly lower than the 70

% accuracy reported in Geisler et al. When training the model with only the most relevant features of the classification in the recovered sample from Geirsler et al. in a sample of first-onset patients with AN, the classification accuracy significantly decreases. This indicates that the model does not generalize, and the extracted features and patterns do not extend to the whole AN population.

In general, it is challenging to ascertain the function or role of a node in such a classification system. The features used by the SVM represent local connectivity metrics (local metrics) of region within the brain (nodes). The local metrics that underpin the classification are the participation index (PI), degree, clustering coefficient (CCloc) and normalized local efficiency (LEGE) of the node in question. The local metric that is most often selected is the PI, which is a metric that quantifies the diversity of connections of a given node with distinct communities. This suggests AN results in alterations that may be related to the communication between distinct modules of the brain. Another metric that is often used by the SVM (4) is the degree of a node which is defined as the number of adjacent edges associated with that node. It is possible that alterations in a node's degree may be related to the importance of the node within a given community. A reduction in degree has been previously observed in individuals with AN, particularly in nodes of the insula (Collantoni et al. 2021). Furthermore, this has been identified as a key feature in the classification of recovered AN patients from TD (Geisler et al. 2020). The local efficiency of a node is a metric that demonstrates the nodes's capacity to communicate with the other nodes within a community. Altered LEGE has also been documented in previous papers that look into differences in graph metrics in the AN population, with increased efficiency observed in the posterior occipital region (Geisler et al. 2016), which is a region present as relevant in our results, but associated with CCloc (feature 3). However, the findings were not robust. Finally, the CCloc is a measure of segregation and alteration in this metric could reflect differences in the ability of a node to integrate a sub-network Geisler et al. 2020.

The results of our study indicate that classification is primarily driven by regions of the post-cingulate, insula, temporal cortex, parietal cortex and cerebellum, with nodes from postoccipital and precuneus also present. Previous classifications using SVM on AN have similarly identified nodes in temporo-parietal brain regions as key contributors to case-control classification (Gaudio and Olivo, 2018).

The insula (features 1, 5 and 18) plays a pivotal role in integrating cognitive control of internal and external multi-sensory stimuli (Molnar-Szakacs and Uddin, 2022). Additionally, there have been reports of alterations in its functional activity in cases of AN. In addition, both the posterior cingulate cortex (features 10 and 16 and is a key member of the limbic system) and the cerebellum coordinate a wide range of functions and processes, including sensorimotor coordination and a variety of cognitive, affective, social, and behavioral processes (De Benedictis et al., 2022). Both of these regions have been found to have altered connectivity in studies on AN (Gaudio et al. 2016).

There is minimal spatial overlap with Geisler et al. 2020 work in the specific nodes that have the 10 most positive or negative weights, which is to be expected given that the nodes in both papers are relatively small (10 mm diameter spheres). However, even though there is low spatial overlap, since several nodes can represent the same networks and regions, there is some overlap in networks and regions related to the prediction of AN. Furthermore, the coefficients of the linear SVM may not be the optimal metrics for interpreting the relationship between a behavioral feature and its function in the brain, whether in terms of its placement in the brain or its role in brain topology. This may translate into limited generalizability of the features identified in recovered anorexia nervosa patients to a first-onset population.

Although we were able to obtain a modest classification accuracy, this work presents itself with several limitations and suggestions for future analysis. First, the features used, including the selected nodes and the graph metrics, may not be optimal. This is because the nodes were not selected based on other literature specific of AN, and no feature reduction step was applied. Second, the extensive number of features used in the model and the characteristics of the selected classifier render it challenging to interpret the results in terms of brain function and their relation to the key features of AN, using this particular method. Future

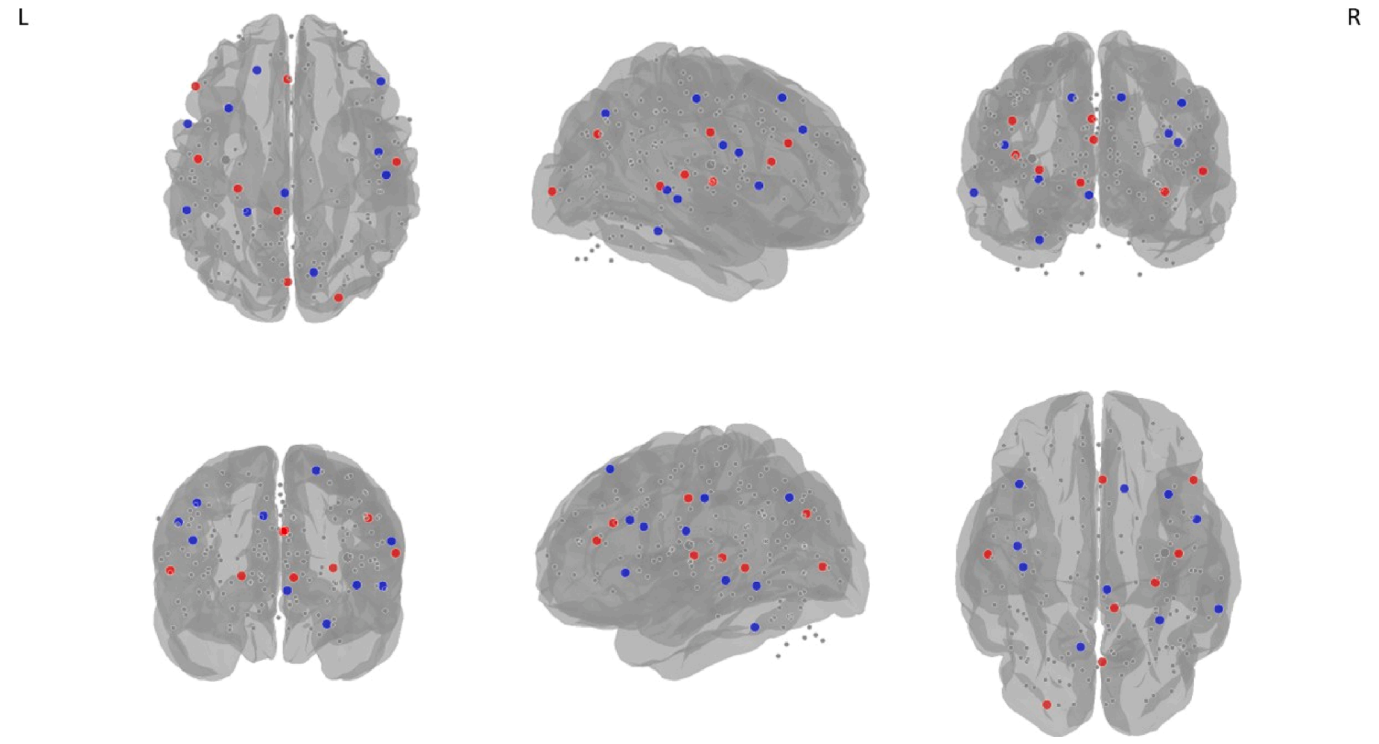


Fig. 3. Location of nodes corresponding to the SVM features with the most positive and most negative values. The 10 most positive weights are represented in red, and the 10 most negative in blue. The corresponding feature names and weights are listed in Table 4.

work should focus on developing methods that not only yield accurate prediction results but also provide better interpretable insights, into the function of each region in AN, using, for instance SHAP values, or interpretable deep learning methods Munroe et al. 2024. With regard to the application of machine learning and rs-fMRI in the prediction of AN based on the detection of a biomarker, further studies would be necessary to ascertain whether these patterns can be identified and linked to the key features of AN, and above all reproduces across cohorts. This could facilitate the early detection of the disorder in a clinical setting, since an accuracy of 0.64 is not enough for detection in a practical setting.

AN has been shown to be associated with comorbid psychiatric conditions (Neale and Hudson, 2020). We found higher rates of both depression and anxiety compared to the control population. In addition, both depression and anxiety have also been associated with differences in functional connectivity using fMRI studies (Li 2022; Xu et al. 2019. While these comorbid conditions in AN are not necessarily independent (i.e., symptoms of malnourishment can mimic depressive symptoms), it is possible that a part of our prediction is a result of identifying the functional connectivity patterns associated with these comorbid conditions. More strict inclusion criteria will be needed to address this question.

Another limitation of the study, which is common to most rs-fMRI and machine learning methods, is the sample size. Overall, studies have reported that the appropriate sample size depends mainly on the supervised model that is used and its complexity (Khosla et al. 2019). Furthermore, data collection has taken place to be harmonized to the Generation R Study (White and Muetzel, 2018), which is a large population based study of child development, therefore, an older sequence is used that allowed harmonization with our epidemiologic studies of child development. Future research could concentrate not only on feature extraction and decision-making processes based on the available sampling size for the task of interest but also on the generalisability of the model (Khosla et al. 2019).

Overall, this study provides an exploratory investigation of the underlying functional changes in AN, by reproducing an existing method, further accentuating the importance of testing the generalizability of existing models in the field of neuroimaging and machine learning. Although our prediction was modest, identifying predictors of AN and understanding the neurological systems behind this disorder, continues to be a priority in the field, due to the difficulty for clinicians to predict the course of AN. Multimodal machine learning approaches and deep learning could lead to better models and possibly yield even better neurophenotypes of psychiatric disorders Khosla et al. 2019; White 2022.

Conclusion

In summary, our work applies an SVM method in the classification of AN patients using local graph metrics in a first-onset population of AN, further showing the classification potential of these features. We note there is room for classification accuracy improvement and the interpretability of the most relevant features remains limited, which shows the problem itself is challenging. The model was adapted from an algorithm previously applied in a sample of recovered AN patients. Our findings indicate that, although a modest degree of classification was achieved in our sample, the generalizability of the results between the two approaches, which encompasses both the classification metrics and the features that drive the task, is not optimal. Further work should concentrate on the creation of more interpretable and generalizable models, as well as the application of existing algorithms to different samples of AN. This will allow for an assessment of whether the results remain significant, regardless of acquisition pipeline, disease stage, or profile. This will also facilitate an attempt to identify a link between brain development and the key features of AN that may predispose to the development of other symptoms.

Data and code availability

Request for data will be granted under European Union's General

Data Protection Regulation following an approved data transfer agreement. Per request please contact Gwen Dieleman at g.dieleman@erasmusmc.nl. Code is freely available and can be accessed via the following link: https://github.com/laura-dias29/Brain-connectivity-and-Machine-Learning-Anorexia-NervosaAny questions about the code please contact Laura Monteiro Rente Dias at l.monteirorentedias@erasmusmc.nl

Funding

This study was granted by the Sophia Foundation for Scientific Research (SSWO) (Grant numbers: S15–13, S22–65). The work of TW was supported by the Intramural Research Program from the National Institute of Mental Health (ZIA MH002986).

CRediT authorship contribution statement

Laura Monteiro Rente Dias: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Hugo Schnack:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Daniel Geisler:** Writing – review & editing, Software, Methodology, Formal analysis. **Marcel Reinders:** Writing – review & editing, Supervision, Methodology. **Tonya White:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Gwen Dieleman:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Xucong Zhang:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgements

The authors would like to thank all the participants and their parents. The BRAVE study was conducted within the Erasmus MC-Sophia in close collaboration with Altrecht-Rintveld, Bravis Hos-pital, Curium-LUMC, Elisabeth-TweeSteden hospital, Emergis, Emergis-Ithaka, GGNet-Amarum, GGZ Delfland, GGZ-Rivierduinen, GGZ-WNB, Franciscus Gasthuis and Vlietland hospital, Reinier de Graaf Gasthuis, Stichting Human Concern and the Van Weel Bethesda hospital. We would like to thank all our collaboration partners for introducing the BRAVE study to the AN participants.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.pscychresns.2026.112255.

References

American Psychiatric Association, DSMTF, DS American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental disorders: DSM-5*. Vol. 5. 5. American psychiatric association, Washington, DC.

Arbabshirani, Mohammad R., et al., 2017. Single subject prediction of brain disorders in neuroimag-ing: Promises and pitfalls. Neuroimage 145, 137–165.

Boehm, Ilka, Geisler, Daniel, King, Joseph A., et al., 2014, October. Increased resting state functional connectivity in the fronto-parietal and default mode network in anorexia nervosa. Front. Behav. Neurosci. 8, 346.

Boehm, Ilka, Geisler, Daniel, Tam, Friederike, et al., 2016. Partially restored resting-state func-tional connectivity in women recovered from anorexia nervosa. J. Psychiatry Neurosci. 41 (6), 377–385.

Collantoni, Enrico, et al., 2021. Brain networks in eating disorders: A systematic review of graph theory studies. Eat. Weight Disord.-Stud. Anorexia Bulimia Obesity 1–15.

Cowdrey, Felicity A., et al., 2014. Increased resting state functional connectivity in the default mode network in recovered anorexia nervosa. Hum. Brain Mapp. 35 (2), 483–491.

De Benedictis, Alessandro, et al., 2022, February. Networking of the Human cerebellum: From anatomo-functional development to neurosurgical implications. Front. Neurol. 13.

Gaudio, Santino, Olivo, Gaia, et al., 2018. Altered cerebellar–insular–parietal–cingular subnetwork in adolescents in the earliest stages of anorexia nervosa: A network–based statistic analysis. Transl. Psychiatry 8 (1), 127.

Gaudio, Santino, Piervincenzi, Claudia, et al., 2015, June. Altered resting state functional connec-tivity of anterior cingulate cortex in drug naive adolescents at the earliest stages of anorexia nervosa. Sci. Rep. 5 (1), 1–11.

Gaudio, Santino, Wiemerslage, Lyle, et al., 2016. A systematic review of resting-state functional-MRI studies in anorexia nervosa: Evidence for functional connectivity impairment in cognitive control and visuospatial and body-signal integration. Neurosci. Biobehav. Rev. 71, 578–589.

Geisler, Daniel, Borchardt, Viola, Boehm, Ilka, et al., 2020. Altered global brain network topology as a trait marker in patients with anorexia nervosa. Psychol. Med. 50 (1), 107–115.

Geisler, Daniel, Borchardt, Viola, Lord, Anton R., et al., 2016. Abnormal functional global and local brain connectivity in female patients with anorexia nervosa. J. Psychiatry Neurosci. 41 (1), 6–15.

Ghosh, Sourish, Dasgupta, Anasuya, Swetapadma, Aleena, 2019. A study on support vector machine based linear and non-linear pattern classification. In: 2019 International Conference on Intelligent Sustainable Systems (ICISS). IEEE, pp. 24–28.

Haynos, Ann F., et al., 2019. Resting state functional connectivity of networks associated with reward and habit in anorexia nervosa. Hum. Brain Mapp. 40 (2), 652–662.

Kaye, Walter H., Fudge, Julie L., Paulus, Martin, 2009. New insights into symptoms and neurocircuit function of anorexia nervosa. Nat. Rev. Neurosci. 10 (8), 573–584.

Kaye, Walter H., Christina E, Wierenga, et al., 2013. "Nothing tastes as good as skinny feels the neurobiology of anorexia nervosa. Trends Neurosci. 36 (2), 110–120.

Khosla, Meenakshi, et al., 2019. Machine learning in resting-state fMRI analysis. Magn. Reson. ImAging 64, 101–121.

King, Joseph A., et al., 2018. Structural neuroimaging of anorexia nervosa: Future directions in the quest for mechanisms underlying dynamic alterations. Biol. Psychiatry 83 (3), 224–234.

Lee, Megan H., Smyser, Christopher D., Shimony, Joshua S., 2013. Resting-State fMRI: A review of methods and clinical applications. Am. J. Neuroradiol. 34 (10), 1866–1872.

Li, Xue-Ying, Chen, Xiao, Yan, Chao-Gan, 2022. Altered cerebral activities and functional connectivity in depression: A systematic review of fMRI studies. Quant. Biol. 10 (4), 366–380.

Molnar-Szakacs, Istvan, Uddin, Lucina Q., 2022. Anterior insula as a gatekeeper of executive control. Neurosci. Biobehav. Rev. 139, 104736.

Munroe, Lindsay, et al., 2024. Applications of interpretable deep learning in neuroimaging: A comprehensive review. Imag. Neurosci. 2, 1–37.

Neale, Josephine, Hudson, Lee D., 2020. Anorexia nervosa in adolescents. Br. J. Hosp. Med. 81 (6), 1–8.

Papadopoulos, Fotios C., et al., 2009. Excess mortality, causes of death and prognostic factors in anorexia nervosa. Br. J. Psychiatry 194 (1), 10–17.

Pereira, Francisco, Mitchell, Tom, Botvinick, Matthew, 2009. Machine learning classifiers and fMRI: A tutorial overview. Neuroimage 45 (1), S199–S209.

Scaife, Jessica Clare, et al., 2017, March. Reduced resting-State functional connectivity in current and recovered restrictive anorexia nervosa. Front. Psychiatry 8, 30.

Schmidt, Ulrike, et al., 2016. Eating disorders: The big issue. Lancet Psychiatry 3 (4), 313–315.

Steegers, Cathelijne, et al., 2024. Brain, behavior, cognition, and physical health in first-onset adolescent anorexia nervosa: The BRAVE Study design and cohort profile. Apert. Neuro 4.

Stevens, Michael C., 2016. The contributions of resting state and task-based functional connectiv-ity studies to our understanding of adolescent brain network maturation. Neurosci. Biobehav. Rev. 70, 13–32.

Van Eeden, Annelies E., Hoeken, Daphne Van, Hoek, Hans W., 2021. Incidence, prevalence and mortality of anorexia nervosa and bulimia nervosa. In: Curr. Opin. Psychiatry 34 (6), 515–524.

Walton, Esther, et al., 2022. Brain structure in acutely underweight and partially weight-restored individuals with anorexia nervosa: A coordinated analysis by the ENIGMA Eating Disorders Working Group. Biol. Psychiatry 92 (9), 730–738.

Wen, Junhao, et al., 2024. Dimensional neuroimaging endophenotypes: Neurobiological represen-tations of disease heterogeneity through machine learning. Biol. Psychiatry 96, 564–584.

White, Tonya, 2022. Behavioral phenotypes, stochastic processes, entropy, evolution, and individ-ual variability: Toward a unified field theory for neurodevelopment and psychopathology. In: Aperture Neuro 1–3.

White, T., Muetzel, R., et al., 2014. Time of acquisition and network stability in pediatric resting-state functional magnetic resonance imaging. Brain connect. 4 (6), 417–427.

White, T., Muetzel, R.L., et al., 2018. Paediatric population neuroimaging and the Generation R Study: the second wave. Eur. J. Epidemiol. 33 (1), 99–125.

Xu, Jie, et al., 2019. Anxious brain networks: A coordinate-based activation likelihood estimation meta-analysis of resting-state functional connectivity studies in anxiety. Neurosci. Biobehav. Rev. 96, 21–30.