Machine learning based classification of excessive smartphone users via neuronal cue reactivity

Jailan Oweda [a,b,*], Mike Michael Schmitgen [a], Gudrun M. Henemann [a], Marius Gerdes [b], Robert Christian Wolf [a]

[a] Department of General Psychiatry, Heidelberg University Hospital, Germany
[b] Karlsruhe Institute of Technology, Germany

ABSTRACT

Excessive Smartphone Use (ESU) poses a significant challenge in contemporary society, yet its recognition as a distinct disorder remains ambiguous. This study aims to address this gap by leveraging functional magnetic resonance imaging (fMRI) data and machine learning techniques to classify ESU and non-excessive smartphone users (n-ESU) based on their neural Cue-Reactivity (CR) signatures. By conducting a CR task and analyzing brain activation patterns, we identified spatial similarities between addictive smartphone use and established addictive disorders. Our approach involved employing Support Vector Machines (SVM) for classification, enhanced with feature selection methods such as Recursive Feature Elimination (RFE) and Model-based Selection and dimensionality reduction methods such as and Principal Component Analysis (PCA) to mitigate the challenges posed by limited dataset size and high dimensionality of fMRI data. The results demonstrate the effectiveness of our classification model, achieving accuracies of up to 79.9 %. Furthermore, we observed region-specific activations contributing significantly to classification accuracy, highlighting the potential biomarkers associated with ESU. External validation on longitudinal data revealed the necessity for larger training datasets to improve model generalizability. Additionally, feature selection techniques proved crucial for optimizing model performance, particularly in datasets with combined information from multiple sources. Our findings underscore the importance of incorporating more data to enhance model stability and generalizability, with implications for advancing the understanding and treatment of ESU and related disorders. Overall, our study demonstrates the promise of machine learning approaches in elucidating neural correlates of ESU and informing targeted interventions for affected individuals.

1. Introduction

Excessive Smartphone Use (ESU), sometimes also referred to as "smartphone addiction", remains debatable and is not yet recognized by the DSM-5 (DSM-5TR). The main reason is that smartphones can be used for various purposes and thus include several addiction types, including Internet Gaming Disorder (IGD) and related syndromes, e.g. Internet Addiction Disorder (IAD) (Lawrence Robinson and Jeanne Segal 2023). However tools have been created that can estimate "excessiveness" of smartphone usage. For example, the Smartphone Addiction Inventory (SPAI) (Lin et al. 2014) is a 26-item self-report tool designed to assess problematic smartphone use based on criteria like compulsive behavior, tolerance, withdrawal, and negative impacts on health and social functioning. The SPAI, developed with reference to excessive

smartphone use and Internet Gaming Disorder (IGD), has shown strong internal consistency and reliable test-retest performance in previous studies (Pavia et al. 2016; Lin et al. 2014). Also, the Smartphone Addiction Scale-Short Version (SAS-SV) is measured through a self-reported questionnaire and assesses different domains such as daily-life disturbances, withdrawal symptoms, tolerance, and virtual social relationships, which are affected by excessive smartphone use (Hamamura et al. 2023). Additionally, tracking the number of hours spent on smartphones daily can serve as a further indicator of excessive use. In this context, "nonexcessive users" refers to average smartphone users, who may not exhibit problematic behaviors but still engage in regular daily smartphone use.

In a recent study (Mike M. Schmitgen et al. 2020), an fMRI experiment was conducted to compare ESU and non-excessive smartphone

users (n-ESU) during a Cue reactivity (CR)-task, which consisted of stimuli that were either pictures of smartphones (turned on or off) or neutral pictures (Fig. 1). Significant CR-related activity was observed in the medial prefrontal cortex (mPFC), occipital cortex, temporal cortex, Anterior Cingulate Cortex (ACC), temporoparietal regions, and cerebellum when contrasting images of smartphones vs. neutral stimuli (Mike M. Schmitgen et al. 2020). Additionally, for the contrast between active vs. inactive smartphones, differences were found in the frontal operculum/anterior insula and precentral gyrus. These results indicate that there are spatial similarities in cue-reactivity-related brain activation between addictive smartphone use and other well-known addictive disorders (Mike M. Schmitgen et al. 2020). It is unknown so far, whether such activity patterns could be used for classification purposes, i.e. for approaches that seek to delineate distinct patterns of neural activity that could reveal group-specific biological signatures. Such signatures could decisively inform biological models of ESU and related conditions, such as IGD or other technology-related addictive behavior.

This paper aims to train machine-learning models that classify the ESU and n-ESU subjects based on their neural Cue-Reactivity signature and to test the model on a separate dataset. For this purpose task-based functional MRI (fMRI) is used, which captures brain activity by measuring bloodoxygen-level-dependent (BOLD) signals during specific tasks.

Preprocessing steps are essential to prepare fMRI data for analysis. These typically include Slice Time Correction (Adjusts for differences in the time it takes to acquire each slice in a volume, ensuring temporal alignment of the data), Motion Correction and Realignment (Corrects for subject movement, ensuring consistent spatial alignment across images), Segmentation (Separates brain tissues (gray matter, white matter, cerebrospinal fluid) for more precise analysis), Normalization (Aligns individual brains to a standard template for easier group-level comparisons), and Smoothing (Applies a spatial filter to reduce noise and enhance signal detection).

After preprocessing, traditional fMRI analysis usually involves two stages First-Level and Second-Level Analysis. In First-Level Analysis of task-based fMRI, models are applied to each participant's data to create contrast images, which show differences in brain activity between different conditions. These contrast images are then used in Second-Level Analysis to make comparisons across a group of participants. However, machine learning can replace Second-Level Analysis by looking for patterns directly across participants' data. Instead of averaging, algorithms find subtle patterns and make predictions for individual cases, capturing more complex in- formation and providing deeper insights.

A common issue with fMRI data arises from the limited database size, primarily due to the high cost associated with its acquisition. To address this challenge, data augmentation is often employed to expand the dataset by applying random transformations to existing samples and

Fig. 1. Schematic overview of the cue reactivity task (see supplementals for image examples from each group).

generating new images. While traditional 2D scans can be manipulated through spatial transformations like rotation and scaling, fMRI data is 4-dimensional (3D + time), making data augmentation more complex (Zhuang et al., 2019). The distinctive characteristics of fMRI, along with the high correlation between neighboring voxels and time steps, pose challenges for standard data augmentation techniques (Ghassemi et al., 2020). These techniques may struggle to capture spatial and temporal dependencies adequately, potentially resulting in unrealistic or inappropriate variations in the data. Commonly Generative Adversarial Networks can be used for realistic augmentations, however, as this is a deep learning method it already requires a high number of samples (Ghassemi et al., 2020; Zhuang, Schwing, and Koyejo 2019).

When faced with smaller datasets, it is conventional to use simpler machine learning approaches that can perform well with fewer samples and avoid overfitting. The choices include:

(a) Support Vector Machines (SVM) work by finding a hyperplane that maximizes the margin, which refers to the maximum width of the boundary that separates the different classes, as it tends to generalize better on unseen data. The support vectors, which are data points closest to the decision boundary, control the margin width. These models are well-suited to high-dimensional data because they can handle complex, non-linear relationships using kernel functions (Steinwart et al. 2008; Abe 2010).

(b) Random Forests (RF) is an ensemble learning method that combines multiple decision trees for improved accuracy. This approach reduces variance and is more robust to overfit- ting and can handle noisy data providing more reliable predictions and making them more suitable for datasets, where the number of features often exceeds the number of samples. However, it is crucial to apply hyper-tuning for example to the number of trees and their depths for better performance (Kamarajan et al., 2020; Afis et al., 2024).

(c) Multilayer Perceptrons (MLP) are a type of neural network with one or more hidden layers. While they are more complex than SVMs and Random Forests, they can still be effective with smaller fMRI datasets, especially when carefully regularized. They learn complex representations by transforming input features through nonlinear activation functions. Their success depends on the availability of labeled data, model architecture, and training strategies, which is why they need careful hypertuning (Unzueta 2023; Benoit Liquet and Nazarathy 2023; Afis et al., 2024).

Another approach to overcome the limited database obstacle, as well as its high dimensionality, sparsity, and noise is applying feature selection or dimensionality reduction. To test this theory the following methods were chosen as an extension to the previously described machine learning models. As described further in Section 2.8 the ML models were trained with and without these methods to highlight their effects.

(d) Recursive Feature Elimination (RFE) starts by training a model on the entire set of features and ranking the features based on their importance scores obtained from the model's coefficients. Then, the least important features are removed from the dataset. The process is repeated iteratively until the desired number of remaining features (in this case chosen empirically) is reached. It is particularly effective because it considers the interactions between features and captures their combined predictive power (Ranjan and Singh 2023).

(e) "Select From Model" also relies on feature importances, selecting the most important features according to a specified threshold and discarding the rest. This approach is useful when the model's coefficients or feature importances provide insights into which features are most influential for prediction. Both of these methods can be used in combination with Support Vector Machines (SVM) and Random Forests (RF) as these models return feature coefficients.
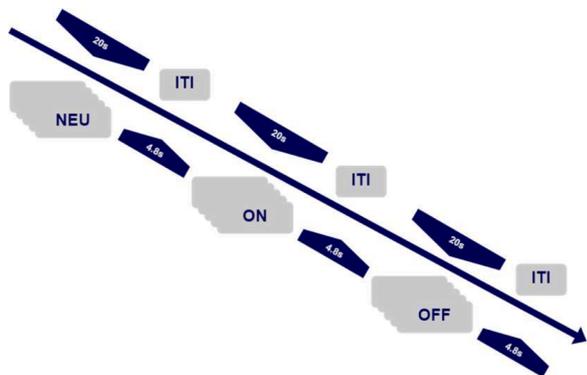
(f) Principal Component Analysis (PCA) is a dimensionality reduction technique used to transform high-dimensional data into a lower-dimensional space while preserving most of the variance in the data. This model-independent method works by identifying the principal components, which are linear combinations of the original features that capture the maximum variance. These principal components are orthogonal to each other, meaning they are uncorrelated. By retaining only a subset of the principal components that explain most of the variance, PCA reduces the dimensionality of the data while minimizing information loss (Xie et al. 2009).

We predicted that ESU will exhibit distinct neural activation patterns during the cr-task compared to n-ESU and that these patterns will be detectable and classifiable using machine-learning models. We also expected that feature selection and dimensionality reduction techniques, such as RFE, Model-based Selection, and PCA, will enhance the classification accuracy of the machine-learning models by identifying the most relevant neural features from the fMRI data.

Here, we trained machine-learning models on regional activations extracted from fMRI data from a CR-task to predict ESU or n-ESU state in a sample of young adult smartphone users. We tested SVM, MLP, RF. As SVM demonstrated the best performance, it was chosen for further analyses in this study. MLP and RF results were added to the supplemental material.

2. Methods

2.1. Participants

Participants in this study were subsets of previous CrossSectional studies by Horvath et al. (2020) and Schmitgen et al. (2020) and an ongoing, longitudinal study of the same group. The two datasets consist of completely different in- dividuals. Their recruitment was conducted through flyers, posters at Heidelberg University, the city center, and social media advertisements. Two user groups, excessive smart- phone users (ESU) and controls (n-ESU), were defined based on the Smartphone Addiction Scale-Short Version (SAS-SV) (Hamamura et al. 2023) with cutoff values of >31 for males and >33 for females resulting in 20 ESU and 22 n-ESU in the cross-sectional dataset and 19 ESU and 17 n-ESU in the longitudinal dataset.

Before the MRI scans were acquired, participants com- pleted assessments including the Smartphone Addiction Inventory (SPAI) (Lin et al. 2014), Beck Depression Inventory (BDI)-II (Gellman and Turner 2013), and Barratt Impulsive- ness Scale version 11 (BIS-11) (Barratt 1975). These mea- sures showed satisfactory reliability and validity for assessing smartphone addiction, depression, and impulsivity, respectively (Hamamura et al. 2023; Lin et al. 2014; Gellman and Turner 2013).

Participants were required to abstain from smartphone use during psychometric and MRI assessments. In the longitudinal study, they were also required to abstain from smartphone use for 72 h before running a second MRI acquisition. Approved by the Ethics Committee of the Medical Faculty at Heidelberg University, the study adhered to the Declaration of Helsinki. All participants provided written informed consent and received monetary compensation for their participation.

2.2. Subject demographics

Participants with poor data quality were excluded after visually inspecting the fMRI scans for artifacts, such as signal intensity distortions and inconsistencies throughout the entire time series. Additionally, subjects with head movements exceeding 3 mm or 3° were excluded from the analysis.

This left us with 42 subjects in the cross-sectional dataset (20 Excessive Smartphone Users (ESU), 22 from the control group (HC)). 30 of them were female and 12 were male. Their mean age was 22.8 3.14 std, the youngest being 18 and the oldest 30 years old. The mean Smartphone Addiction Inventory (SPAI) -Score of the ESU group was 56.95 10.2 std, with a minimum score of 38 and a maximum of 82. The mean SPAI-Score of the control group was 36 6.9, with a minimum score of 27 and a maximum of 55.

In the Longitudinal dataset, 36 subjects were kept after data quality inspection (19 ESU, 17 HC). 20 of them were female and 16 were male. Their mean age was 22.7 1.93 std, the youngest being 18 and the oldest 26 years old. The mean SPAI-Score of the ESU group was 59.2 13.2 std, with a minimum score of 35 and a maximum of 90. The mean SPAI-Score of the control group was 40 10.47, with a minimum score of 28 and a maximum of 66.

The cross-sectional and longitudinal datasets contained different subjects; in both datasets, there was an ESU and an HC group. Subsequently, during machine learning training the major goal was to classify these 2 groups using several distinct approaches. We used one dataset and then mixed both of them to compare the results. In all cases, we split the data to separate the test data from the training data strictly.

2.3. Data acquisition

A 3-T Magnetom TIM Trio MR Scanner manufactured by Siemens in Erlangen, equipped with a 32-channel head coil, was utilized in the cross-sectional study to acquire comprehensive whole-brain structural and functional scans in a dimly lit environment. For the longitudinal study a 3-T Magnetom Prisma Fit Scanner, also manufactured by Siemens, was used. Participants' heads were securely fixed in the head coil using foam cushions.

The scanning protocol of the cross-sectional study comprised four sequential, functional measurements, specifically a resting-state scan, three experimental paradigms, and a structural scan, as Horvath et al. 2020 and Mike M. Schmitgen et al. 2020 outlined. In the longitudinal study, the same acquisitions were reacquired after the participants abstained from their phones for 72 h. The results of the different experiments and their combinations were already reported in several studies (Mike M. Schmitgen et al. 2020; Horvath et al. 2020; Mike M Schmitgen et al. 2022; Hirjak et al. 2022; Henemann, Mike M Schmitgen, Wolf, Hirjak, Kubera, Sambataro, Bach, et al. 2023; Henemann, Mike M Schmitgen, Wolf, Hirjak, Kubera, Sambataro, Lemenager, et al. 2023).

2.4. CR-task

A modified CR task, adapted from Beck et al. 2012, was employed to examine cue-specific brain activation in both datasets. The task involved the presentation of images depicting neutral, non-modern, non--media-related stimuli such as furniture, plants, landscapes, and animals without humans (NEU). Additionally, images included smartphones in a nonoperating state (OFF) and smartphones in use (ON) displaying main or dialing screens, or apps from various categories. The stimuli were standardized in size (1024 × 768 pixels) but were not matched for valence or physical image properties. The order of condition blocks and pictures within the blocks was randomized across subjects. Images from each condition were presented in blocks of five (20 s per block), separated by a 4.8-second presentation of a fixation cross (ITI) without jittering in the interstimulus intervals (see Fig. 1) (Beck et al. 2012).

2.5. Data preprocessing

Imaging data were pre-processed using the Nipype (Nipype: Neuro-imaging in Python Pipelines and Interfaces, 2023) module in Python (https://www.python.org/) and more specifically the SPM12 (SPM12: Statistical Parametric Mapping, 2023) interface package. Nipype was used to create a pipeline that realigned the functional data, coregistered it with structural images, segmented for normalization to standard MNI space, and smoothed with a 9-mm Gaussian kernel at full width at half

maximum (FWHM). Subjects with head motion $>(3 \times 3 \times 3)$ mm or $3°$ were excluded from further analysis.

The study utilized a general linear model (GLM) to detect blood oxygen level-dependent (BOLD) activation. Four regressors were defined: ITI, NEU, and images with phones with screens turned on (ON) and phones with screens turned off (OFF). These regressors were created by convolving the timing of these stimuli with a standard hemodynamic response function (HRF). Additionally, six realignment parameters were included as nuisance regressors. A high-pass filter (128 Hz) was applied to eliminate low-frequency signal drift. Con- trast images comparing neutral and phone images (phone (ON+ OFF) >NEU) and also images turned on and off (phone ON >phone OFF) were generated to investigate cue-induced brain activation. These maps were then utilized to classify the subjects.

Before passing to the classification task, region-specific activations were extracted from the contrast images by masking them with the Neuromorphometrics atlas (Neuromorphometrics,Inc. - Building a Model of the Living Human Brain, 2023) using the Nilearn (Nilearn 2024) module in Python, resulting in 125 features per subject per contrast (136 before removing masks not including grey matter, such as ventricles and cerebrospinal fluid). In this case, two contrasts are used for each subject: 'Phone (ON+OFF) > NEU' and 'Phone ON > Phone OFF' resulting in 250 features per subject.

2.6. Grid search

Region-specific activations were used to train a Support Vector Machine model that classifies ESU and n-ESU subjects (Random Forests and Multi-layer Perceptrons were also trained but SVM proved to be the most efficient, the results can be found in the supplemental material). The model was developed in Python (2023) using the following modules: Scikit-learn - Machine Learning in Python (2023), Keras (2023) and Tensorflow (2024). The model was hyper-tuned using a k-fold ($k = 10$) cross-validation method to find the parameters that deliver the best accuracy during validation. The k-folds are repeated n times ($n = 5$) while shuffling the data to guarantee the sustainability of the obtained accuracies. The hyperparam- eter was kernel type which represents the hyperplane function and the regularization parameter (C), which is a measure of the tolerance for misclassification allowed in the model.

The model was trained and tested three times using: (i) The cross-sectional dataset (42 subjects (20 ESU, 22 HC (nESU)), (see supplementals for further demographics about each class)) for both training and testing (with cross-validation). (ii) The cross-sectional dataset for training and the longitudinal (36 subjects (19 ESU, 17 HC)) for testing (in the supplemental material) (iii) A combination of both datasets for training and testing (cross-validation). In all three combinations, we ensured that training and test data were strictly separated. Testing was conducted exclusively on subjects not part of the training sample. For the splitting the stratified method was used to preserve the proportion of ESU and HC in both training and test date.

2.7. Choosing the ML model

Using grid search, we optimized the hyperparameters for three different models: SVM, Random Forest, and MLP. Each optimized model was then trained on the three train/test combinations described in Section 2.6 and evaluated as outlined in Section 2.9.

SVM outperformed the other two models, particularly after applying feature selection, as detailed in Section 2.8. Due to this, we chose to focus on the SVM for the results and discussion, since discussing all models would overwhelm the manuscript and appear redundant, especially when some models, such as those with an accuracy of around 50 %, indicated random classification. However, the results from RF and MLP are included in the supplemental material for reference.

2.8. Feature selection

As the used dataset contains 125 features and only 42 samples in one dataset and 36 in the other, the model's performance has been compared to when RFE or Select from Model methods are applied for feature selection or when PCA is applied for dimensionality reduction.

To reduce the risk of overfitting, the number of data points should be at least 10 times the number of features (Smolic 2024). This is why we chose to compare the models with 4 (since one dataset includes 40 subjects), 8 (since the 2 datasets combined include 80 subjects), and 16 (to verify whether the increase of features will lead to overfitting) selected features.

2.9. Evaluating ML models

In addition to accuracy, which is the number of correct predictions over the total number of predictions other scores have also been evaluated. Sensitivity or Recall (*Re*) is the proportion of ESU who test positive: $P = \frac{\text{True Positives}}{\text{True Positives+False Negatives}}$. Precision (Pr) is the proportion of correctly identified ESU $P = \frac{\text{True Positives}}{\text{True Positives+False Positives}}$. The harmonic mean of both metrics can be used to balance them $P = \frac{2Pr*Re}{Pr+Re}$, also called the F1-score. Specificity is the proportion of n-ESU who test negative: $P = \frac{\text{True Negatives}}{\text{True Negatives+False Positives}}$.

Moreover, receiver-operating-characteristic (ROC) analysis and the area under the ROC curve (AUC) were used to evaluate the performance of the classifiers. AUC represents the classification power of a classifier. The values of AUC range from 0 to 1 and larger AUCs indicate better classification abilities.

The regional contributions to the classification model were ranked in descending order, by extracting the feature coefficients from the models.

3. Results

3.1. Single dataset

In our classification task distinguishing between ESU vs. *n*ESU from a single cross-sectional dataset, the optimal model emerged as utilizing an SVM with a linear kernel and a regularization parameter of C = 0.005 (Table 1). This model showcased superior performance when integrated with the Recursive Feature Extractor methodology, specifically employing 16 selected features. This approach yielded a mean accuracy of 79.9 % (18.0 % standard-deviation).

It was noted that specific regions demonstrated greater significance in the classification task, and these regions manifested from different contrast images. Notably, in the Phone (ON + OFF) > NEU contrast (Fig. 2a), the Right Anterior Cingulate Gyrus (ACgG), Right Angular Gyrus (AnG), Right Superior Frontal Gyrus (SFG), and Right Superior Frontal Gyrus Medial Segment (MSFG) were identified as the most significant regions, as they were chosen in >40 % of the cross-validation (CV) iterations. Conversely, in the Phone ON > OFF contrast (Fig. 2a), the Left Frontal Operculum (FO) region emerged as the most significant and was chosen in 30 % of the iterations (Table 3a).

3.2. Two combined datasets

Once two datasets (1 cross-sectional and 1 longitudinal) were combined for cross-validated training and testing, the most effective model emerged employing a regularization parameter C = 0.025 within a linear SVM framework (Table 2). Utilizing the RFE method and selecting 16 features yielded the highest accuracy of 78.9 % (12.8 %).

Here, in the Phone (ON + OFF) > NEU (Fig. 3a) contrast the Right ACgG, Right Cerebellum, Left Posterior Insula (PIns), and Right SFG were highlighted as the most significant with a selection rate higher than 40 %. In the Phone ON > OFF contrast (Fig. 3b) the Right Orbital Part of the Inferior Frontal Gyrus (OrIFG) and Right Planum Polare (PP) held

Table 1
Cross-validated mean (stand-deviation) scores from support vector machine model (regularization parameter: C = 0.005) trained on the cross-sectional dataset with different feature selection or dimensionality reduction methods; green color indicates mean accuracies above 75 %.

Feature Selection / Dimensionality Reduction?	# Features	Accuracy	Precision	Recall	F1-Score	Specificity	AUC
None	*250*	75.8 % (±18.8 %)	74.3 % (±34.0 %)	67.0 % (±34.1 %)	67.6 % (±30.3 %)	83.0 % (±25.3 %)	75.0 % (±19.6 %)
	4	52.8 % (±6.6 %)	4.0 % (±19.6 %)	2.0 % (±9.8 %)	2.7 % (±13.1 %)	100.0 % (±0.0 %)	51.0 % (±4.9 %)
RFE	*8*	74.5 % (±14.7 %)	81.3 % (±38.4 %)	49.0 % (±29.1 %)	59.6 % (±30.2 %)	98.7 % (±6.5 %)	73.8 % (±18.8 %)
	16	79.9 % (±18.0 %)	80.3 % (±34.1 %)	68.0 % (±34.3 %)	70.8 % (±31.0 %)	90.3 % (±22.1 %)	79.2 % (±18.8 %)
	4	52.0 % (±4.0 %)	0.0 % (±0.0 %)	0.0 % (±0.0 %)	0.0 % (±0.0 %)	100.0 % (±0.0 %)	50.0 % (±0.0 %)
Model-Based	*8*	68.6 % (±13.6 %)	66.0 % (±47.4 %)	35.0 % (±26.9 %)	45.3 % (±33.2 %)	100.0 % (±0.0 %)	67.5 % (±13.5 %)
	16	76.8 % (±19.3 %)	80.7 % (±29.1 %)	69.0 % (±29.8 %)	71.4 % (±25.5 %)	84.0 % (±24.9 %)	76.5 % (±19.5 %)
	4	65.3 % (±18.6 %)	57.3 % (±40.8 %)	50.0 % (±37.4 %)	50.4 % (±34.2 %)	79.0 % (±30.7 %)	64.5 % (±19.5 %)
PCA	*8*	69.7 % (±18.3 %)	68.7 % (±36.0 %)	59.0 % (±34.2 %)	60.1 % (±29.9 %)	79.0 % (±27.2 %)	69.0 % (±18.6 %)
	16	73.5 % (±18.1 %)	70.7 % (±38.2 %)	60.0 % (±36.1 %)	61.7 % (±32.8 %)	85.3 % (±27.8 %)	72.7 % (±18.7 %)

(a) Contrast: Phone (ON & OFF) > Neutral.

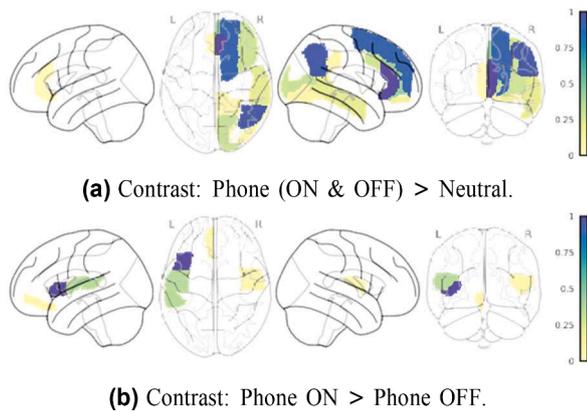**(b)** Contrast: Phone ON > Phone OFF.

Fig. 2. Extracted Regions from the Recursive Feature Elimination (RFE) method combined with the Support Vector Machine model applied on the cross-sectional dataset only. The color is an indicator of the rate of its choice since RFE was applied on 10 folds and repeated 5 times. Purple is a rate up to 100 % while yellow is closer to 0 % (Table 3a).

significance with selection rates over 68 % (Table 3b).

3.3. Features in common

Remarkably, among the 16 regions selected for each of the two training scenarios, four regions were common: Right ACgG, Right SFG, and Right Lingual Gyrus (LiG) from the Phone (ON + OFF) > NEU contrast, along with the Left FO from the Phone ON > OFF contrast.

4. Discussion

In this study, we extracted brain activation during a CR task to classify ESU vs. n-ESU individuals using an SVM model. Three main findings emerged: 1. specific brain regions, such as the Anterior Cingulate Cortex (ACC), Superior Frontal Gyrus (SFG), Lingual Gyrus (LiG), and Frontal Operculum, are significant contributors to ESU classification, 2. Feature selection, especially Recursive Feature Elimination (RFE), played an important role in enhancing model performance and improving its accuracy, 3. Larger training datasets are needed to improve model generalizability and performance.

4.1. Extracted features

The most significant brain region that was agreed upon in both SVM models was the Right ACgG, which is highly related to addiction disorders since it influences various cognitive processes such as decision-making, emotional processing, inhibition, self-regulation, and motivation (Posner et al. 2007). Studies have shown that stimulating the ACC may be effective for treating Substance Use Disorders (SUD) (Zhao et al. 2020).

The following region was the Right SFG, which is associated with goal-directed behavior, impulsivity, reward processing, craving, and the ability to resist urges all of which may lead to persistent substance or behavioral use. Previous studies have demonstrated that the right SFG could serve as a potential biomarker of IGD and provide clues for its diagnosis and treatment especially due to impulsivity which is linked to this specific region (P. Zhang et al. 2023).

The Right Lingual Gyrus is another brain region that was selected by both SVM models. this one is mainly responsible for visual processing and in this context may be associated with directing attention to the visual stimuli presented in the trials. Since this region was chosen in the contrast of Phone vs. Neutral it may point out to a higher reward anticipation resulting from seeing phone images in ESU subjects potentially resulting in craving and cue-induced reactivity. Although this is not a common biomarker in addiction, the lingual Gyrus may be linked to emotional abnormalities such as depressive disorders (M. Zhang et al. 2021).

The left frontal operculum is the final brain region that was selected by both SVM models. It plays a role in selfawareness, verbal expression, and also motor responses which may be linked to addictive behaviors (Darnai et al. 2019).

4.2. External validation

One of the initial observations upon validating the SVM model trained on the cross-sectional dataset externally on the longitudinal dataset was a numerical decrease in performance, nearly resembling random outcomes (see supplementary material). This decline can be attributed to the size of the training dataset, which fails to introduce sufficient characteristics to effectively train the model. The high standard-deviation of the accuracy in the CV scores which reached up to 19.3 % (Table 1, Select From Model 16 Features) highlights the observation that the choice of the split plays a crucial role in the model's

Table 2

Cross-validated mean (stand-deviation) scores from Support Vector Machine Model (regularization parameter: C = 0.025) trained on a combination of the cross-sectional and longitudinal dataset with different feature selection or dimensionality reduction methods; green color indicates mean accuracies above 75 %.

Feature Selection / Dimensionality Reduction?	# Features	Accuracy	Precision	Recall	F1-Score	Specificity	AUC
None	250	64.8 % (±13.2 %)	69.7 % (±22.7 %)	56.5 % (±22.9 %)	59.7 % (±18.6 %)	73.2 % (±19.7 %)	64.8 % (±13.2 %)
	4	74.7 % (±14.1 %)	77.2 % (±16.9 %)	76.3 % (±21.8 %)	74.3 % (±15.3 %)	72.8 % (±22.7 %)	74.6 % (±14.1 %)
RFE	8	75.5 % (±14.2 %)	78.3 % (±17.8 %)	74.8 % (±21.6 %)	74.3 % (±16.4 %)	75.8 % (±19.8 %)	75.3 % (±14.3 %)
	16	78.9 % (±12.8 %)	82.3 % (±16.7 %)	77.0 % (±21.1 %)	77.5 % (±15.4 %)	80.3 % (±18.8 %)	78.7 % (±12.9 %)
	4	61.1 % (±14.1 %)	65.2 % (±28.8 %)	45.8 % (±23.3 %)	51.3 % (±22.2 %)	76.7 % (±20.0 %)	61.3 % (±14.0 %)
Model-Based	8	65.4 % (±16.2 %)	68.2 % (±26.2 %)	58.7 % (±25.2 %)	60.7 % (±22.7 %)	72.7 % (±23.0 %)	65.7 % (±16.2 %)
	16	73.4 % (±13.9 %)	78.9 % (±21.1 %)	67.7 % (±23.4 %)	70.0 % (±18.9 %)	79.3 % (±21.4 %)	73.5 % (±13.8 %)
	4	57.5 % (±15.7 %)	59.5 % (±30.0 %)	46.8 % (±27.8 %)	48.9 % (±23.8 %)	67.5 % (±25.1 %)	57.2 % (±15.7 %)
PCA	8	59.5 % (±14.6 %)	62.9 % (±24.4 %)	57.0 % (±25.9 %)	56.3 % (±19.5 %)	61.7 % (±25.7 %)	59.3 % (±14.7 %)
	16	56.3 % (±16.4 %)	60.0 % (±23.9 %)	55.0 % (±25.5 %)	53.7 % (±19.3 %)	57.2 % (±28.8 %)	56.1 % (±16.5 %)

(a) Contrast: Phone (ON & OFF) > Neutral.

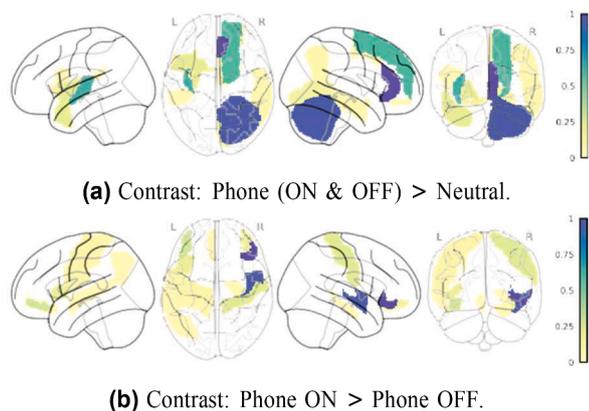**(b)** Contrast: Phone ON > Phone OFF.

Fig. 3. Extracted Regions from the Recursive Feature Elimination (RFE) method combined with the Support Vector Machine model applied on a combination of the cross-sectional and longitudinal datasets. The color is an indicator of the rate of its choice since RFE was applied on 10 folds and repeated 5 times. Purple is a rate up to 100 % while yellow is closer to 0 % (Table 3b).

performance and that more data is needed for better generalizability.

Furthermore, the differences between the experiments may also contribute to the region-specific contrast activations. Specifically, while participants in the first experiment engaged in one-time tasks, those in the second experiment were aware in advance of the requirement to abstain from smartphone usage for 72 h before performing a subsequent round of tasks. This anticipation may have induced excitement or anxiousness, potentially influencing brain functionality during task execution, resulting in different outcomes, than observed in the first experiment.

4.3. Effect of feature selection

For the selected SVMs, it was observed, that when a single dataset was utilized for CV (Table 1), feature selection proved unnecessary, given the already commendable accuracy of 75.8 %. When feature selection was implemented, the model opting for more features numerically consistently exhibited superior performance across all selection methods. This points out that not only a higher number of regions are significant for the classification, but also that the specific combination of the regions plays a role in the task. This can be especially concluded

from the noteworthy improvement of >10 % when 8 features are selected with the RFE method in contrast to when only 4 features are sorted out.

When the model was employed for CV with a fusion of both datasets (Table 2), the necessity for feature selection became more apparent due to a notable drop in accuracy to 64.8 %. This decrease in accuracy points to the difference between activations in both datasets. In this context, RFE emerged as particularly effective, introducing a substantial performance improvement. It resulted in a 10 % enhancement when 4 features were selected and 14 % when 16 features were chosen. Also here, it was noted that whether the RFE or Select From Model method was employed for feature selection, the model performance was better with more features. While the mean accuracy was increasing the variance decreased slightly, too.

4.4. Augmentation through dataset combining

Despite the significant decline in mean accuracy by 11 % when SVM's best model is cross-validated with a combination of two datasets compared to only one dataset without feature selection and the marginal decline by 1 % using the RFE method to select 16 features, an important benefit emerged: a noteworthy reduction of variance by 5.6 % in the first case and 5 % in the second. This reduction suggests a more stable and reliable predictive performance, promising a more robust and accurate classification score overall.

It is crucial to consider that utilizing 16 features with a dataset containing 78 samples enhances generalizability compared to applying the same number of features to a dataset with only 48 samples. This larger sample size allows for a more varied representation of the data distribution, potentially resulting in a more robust and applicable model when deployed in real-world scenarios.

4.5. PCA & regularization parameter

It was observed, that when two datasets were combined the performance when using PCA for dimensionality reduction was significantly lower than without applying this method (by 5.3 % to 8.5 % depending on the chosen number of components). This may point out that the retained principal components do not adequately represent the class separation.

The choice of parameter C for SVM models influences the balance between training error and margin. A smaller value of C allows for a

Table 3

16 Extracted Regions from the Recursive Feature Elimination (RFE) method combined with the Support Vector Machine model. The 2nd and 3rd columns indicate from which contrast image the region has been selected and the rate of its choice since RFE was applied on 10 folds and repeated 5 times. The green color highlights the regions that were in common between case (a) application on a single dataset and (b) application on two combined datasets.

(a) Cross-sectional Dataset only.

Region	Phone > Neutral	Phone ON > Phone OFF
Right ACgG anterior cingulate gyrus	0.96	
Right AnG angular gyrus	0.9	
Right SFG superior frontal gyrus	0.86	
Right MSFG superior frontal gyrus medial segment	0.4	
Left FO frontal operculum		0.3
Right Calc calcarine cortex	0.28	
Right OCP occipital pole	0.26	
Right MFG middle frontal gyrus	0.26	
Right MTG middle temporal gyrus	0.2	
Right AOrG anterior orbital gyrus	0.14	
Left CO central operculum		0.1
Left PT planum temporale		0.1
Right LiG lingual gyrus	0.06	
Right IOG inferior occipital gyrus	0.06	
Right PCgG posterior cingulate gyrus	0.06	
Right CO central operculum		0.02

(b) Cross-sectional and Longitudinal Datasets combined.

Region	Phone > Neutral	Phone ON > Phone OFF
Right OrIFG orbital part of the inferior frontal gyrus		0.72
Right PP planum polare		0.68
Right ACgG anterior cingulate gyrus	0.68	
Right Cerebellum Exterior	0.64	
Left PIns posterior insula	0.42	
Right SFG superior frontal gyrus	0.4	
Left LOrG lateral orbital gyrus		0.18
Right PoG postcentral gyrus		0.16
Left TMP temporal pole	0.14	
Left Putamen	0.14	
Left FO frontal operculum		0.12
Right LOrG lateral orbital gyrus		0.1
Right LiG lingual gyrus	0.08	
Left Basal Forebrain	0.08	
Left AIns anterior insula	0.08	
Left PrG precentral gyrus		0.06

larger margin. It prioritizes a wider separation between classes, even if it means more misclassifications on the training data, and is useful when data points are well-separated and noise/outliers are minimal. On the other hand, a larger value of C emphasizes minimizing the training error. It seeks to fit the training data as accurately as possible, even if it sacrifices margin width, and is beneficial when data points are not well-separated or noise/outliers are significant.

Since the grid search has found a larger value (C = 0.025) to be more performant when two datasets are combined vs. when only one is used (C = 0.005) it indicates, that combining the two datasets has resulted in a more difficult separation of the data and when the observation of the decreased performance with PCA is added to that, it means that the combination has introduced more variability to the dataset avoiding underfitting and making the obtained model more robust and generalizable.

4.6. Limitations

Despite the insights gained, there are several limitations to consider. The relatively small sample size (48 to 78 participants) constrains the generalizability of the findings. Larger and more diverse samples would enhance the robustness and applicability of the results. Additionally,

ensuring consistency in experimental conditions is crucial for more reliable outcomes.

While our approach provides valuable insights, it's important to note that not all available algorithms were tested, and there may be others that could potentially perform better. A comprehensive review of every commonly used method, however, would have been beyond the scope of this report.

5. Conclusions

In conclusion, our data show the ability of simple ML models to classify CR-related fMRI based on extracted regional activations. The effect of applying feature selection methods such as RFE has been highlighted. We show that introducing more data to the model may introduce a significant improvement to the performance, stability, and generalizability of the model. This approach may even be used to find the regions that contribute most to the class separation thus allowing to delineate robust conclusions on neural mechanisms that underlie ESU and related addictive behaviors.

CRediT authorship contribution statement

Jailan Oweda: Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Mike Michael Schmitgen:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization. **Gudrun M. Henemann:** Data curation, Conceptualization. **Marius Gerdes:** Supervision, Methodology. **Robert Christian Wolf:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

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

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.pscychresns.2024.111903.

References

Abe, Shigeo, 2010. Support Vector Machines For Pattern Classification, 2nd ed. Springer London, London. Advances in Pattern RecognitionISBN: 9781849960984. URL: https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=602997.

Afis, Ajala, et al., (2024). Autoencoder-based feature extraction and classification for fMRI-based deep brain stimulation parameter optimization for parkinson textquoterights disease treatment: towards a rapid semi-automated stimulation optimization. In: medRxiv. DOI: 10.1101/2024.01.11.24301179. URL: https://www.medrxiv.org/content/early/2024/01/13/2024.01.11.24301179.

Barratt, Ernest S, 1975. Barratt Impulsiveness Scale. ETS m, 1975.

Ed. by Beck Depression Inventory (BDI), 2013. In: Gellman, Marc D., Turner, J.Rick (Eds.), Encyclopedia of Behavioral Medicine. Springer, New York, pp. 178–179. https://doi.org/10.1007/978-1-4419-1005-9_441. Ed. bySpringer referenceISBN: 978-1-4419-1004-2.

Beck, Anne, et al., 2012. Effect of Brain Structure, Brain Function, and Brain Connectivity on Relapse in Alcohol-Dependent Patients. Arch. Gen. Psychiatry 69 (8), 842–852. https://doi.org/10.1001/archgenpsychiatry.2011.2026. ISSN: 0003-990X.

Benoit Liquet, Sarat Moka and Yoni Nazarathy (2023). *The Mathematical Engineering of Deep Learning (2021) - Chapter 4: general Fully Connected Neural Networks*. URL: https://deeplearningmath.org/general-fully-connected-neural-networks.html (visited on 10/28/2023).

Diagnostic and statistical manual of mental disorders, fifth edition text revision: DSM-5-TR (2022). Fifth edition, text revision. Washington, DC: american Psychiatric Association Publishing. ISBN: 9780890425763.

Darnai, Gergely, et al., 2019. Internet addiction and functional brain networks: task-related fMRI study. Sci. Rep. 9 (1), 15777.

Ghassemi, Navid, Shoeibi, Afshin, Rouhani, Modjtaba, 2020. Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. Biomed Signal Process Control 57, 101678. https://doi.org/10.1016/j.bspc.2019.101678. ISSN: 17468094.

Hamamura, Toshitaka, et al., 2023. Validity, reliability, and correlates of the Smartphone addiction scale–short ver- sion among Japanese adults. BMC Psychol. 11 (1), 78. https://doi.org/10.1186/s40359-023-01095-5. ISSN: 2050-7283.

Henemann, Gudrun M, Schmitgen, Mike M, Wolf, Nadine D, Hirjak, Dusan, Kubera, Katharina M, Sambataro, Fabio, Bach, Patrick, et al., 2023a. Cognitive domain-independent aberrant frontoparietal network strength in individuals with excessive smartphone use. Psychiatry Res. Neuroimaging 329, 111593.

Henemann, Gudrun M, Schmitgen, Mike M, Wolf, Nadine D, Hirjak, Dusan, Kubera, Katharina M, Sambataro, Fabio, Lemenager, Tagrid, et al., 2023b. Neurochemical corre- lates of cue reactivity in individuals with excessive smart-phone use. Eur. Addict. Res. 29 (1), 71–75.

Hirjak, Dusan, et al., 2022. Cortical surface variation in individuals with excessive smartphone use. Develop- Mental Neurobiol. 82 (4), 277–287.

Horvath, Juliane, et al., 2020. Structural and functional cor- relates of smartphone addiction. Addict. Behav. 105, 106334. https://doi.org/10.1016/j.addbeh.2020.106334. URL: https://www.sciencedirect.com/science/article/pii/S0306460319313802.

Kamarajan, C., Ardekani, B.A., Pandey, A.K., Kinreich, S., Pandey, G., Chorlian, D.B., Meyers, J.L., Zhang, J., Bermudez, E., Stimus, A.T., et al., 2020. Random Forest Classification of Alcohol Use Disorder Using fMRI Functional Connectivity, Neuropsychological Functioning, and Impulsivity Measures. Brain Sci. 10, 115. https://doi.org/10.3390/brainsci10020115.

Keras (2023). URL: https://keras.io/(visited on 11/05/2023).

Lawrence Robinson Melinda Smith, M.A. and Ph.D. Jeanne Segal (2023). HelpGuide.org - Smartphone and Internet Addiction. URL: https://www.helpguide.org/articles/addictions/smartphone-addiction.htm (visited on 11/07/2023).

Lin, Yu-Hsuan, et al., 2014. Development and validation of the Smartphone Addiction Inventory (SPAI). PLoS ONE 9 (6), e98312. https://doi.org/10.1371/journal.pone.0098312.

Neuromorphometrics,Inc. - Building a Model of the Living Human Brain (2023). URL: http://www.neuromorphometrics.com/(visited on 10/27/2023).

Nilearn (2024). URL: https://nilearn.github.io/stable/index.html (visited on 02/28/2024).

Nipype: Neuroimaging in Python Pipelines and Interfaces (2023). URL: https://nipype.readthedocs.io/en/latest/(visited on 11/04/2023).

Pavia, Laura, et al., 2016. Smartphone Addiction Inventory (SPAI): psychometric properties and confirmatory factor analysis. Comput Human Behav. 63, 170–178. https://doi.org/10.1016/j.chb.2016.05.039. ISSN: 0747-5632URL: https://www.sciencedirect.com/science/article/pii/S0747563216303661.

Posner, M.I., Rothbart, M.K., Sheese, B.E., et al., 2007. The anterior cingulate gyrus and the mechanism of self-regulation. Cogn. Affect. Behav. Neurosci. 7, 391–395. https://doi.org/10.3758/CABN.7.4.391Posner.

Python (2023). URL: https://www.python.org/ (visited on 11/02/2023).

Ranjan, Ashish, Singh, Vibhav Prakash, 2023. Affirma- tive and negative sentence detection in the brain using SVM-RFE and rotation forest: an fMRI STUDY. SN Comput. Sci. 4 (4). https://doi.org/10.1007/s42979-023-01786-1.

SPM12: Statistical Parametric Mapping (2023). URL: https://www.fil.ion.ucl.ac.uk/spm/software/spm12/(visited on 11/04/2023).

Schmitgen, Mike M., et al., 2020. Neural correlates of cue reactivity in individuals with smartphone addiction. Addict. Behav. 108, 106422. https://doi.org/10.1016/j.addbeh.2020.106422.

Schmitgen, Mike M, et al., 2022. Aberrant intrinsic neural network strength in individuals with "smartphone addiction": an MRI data fusion study. Brain Behav. 12 (9), e2739.

scikit-learn - Machine Learning in Python (2023). URL: https://scikit-learn.org/stable/index.html (visited on 11/05/2023).

Smolic, Hrvoje (2024). How Much Data Is Needed For Machine Learning? URL: https://towardsai.net/p/l/how-much-data-is-needed-for-machine-learning (visited on 08/24/2024).

Steinwart, Ingo, et al., 2008. Support Vector Machines. Information Science and Statistics. Springer Science+Business Media, LLC, New York, NY. https://doi.org/10.1007/978-0-387-77242-4. ISBN: 978-0-387-77241-7.

Tensorflow (2024). URL: https://www.tensorflow.org/(visited on 12/02/2024).

Unzueta, Diego (2023). Fully Connected Layer vs. Convolutional Layer: explained. URL: https://builtin.com/machine-learning/fully-connected-layer (visited on 10/28/2023).

Xie, Song-yun et al. (2009). "Brain fMRI processing and classification based on combination of PCA and SVM". In: pp. 3384–3389. DOI: 10.1109/IJCNN.2009.5179085.

Zhang, Mengzhe, et al., 2021. Shared gray matter alterations in subtypes of addiction: a voxel-wise meta-analysis. Psychopharmacology (Berl.) 238 (9), 2365–2379. https://doi.org/10.1007/s00213-021-05920-w.

Zhang, Pengyu, et al., 2023. Impulsivity-related right superior frontal gyrus as a biomarker of internet gaming disorder. Gen. Psychiatry 36 (4), e100985. https://doi.org/10.1136/gpsych-2022-100985. ISSN: 2517-729X.

Zhao, Yijie, et al., 2020. Anterior cingulate cortex in addiction: new insights for neuromodulation. Neuromodulation: J. Int. Neuromodul. Soc. https://doi.org/10.1111/ner.13291.

Zhuang, P., Schwing, A.G., Koyejo, O., "FMRI Data Augmentation Via Synthesis," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 2019, pp. 1783-1787, DOI: 10.1109/ISBI.2019.8759585.