SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

April 2023 Vol.66 No.4: 835–847
https://doi.org/10.1007/s11427-022-2206-3

Decoding six basic emotions from brain functional connectivity patterns

Chunyu Liu[1], Yingying Wang[2*], Xiaoyue Sun[1], Yizhou Wang[5] & Fang Fang[1,3,4*]

[1]*School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100871, China;*
[2]*Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310028, China;*
[3]*IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China;*
[4]*Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China;*
[5]*Center on Frontiers of Computing Studies, School of Computer Science, Peking University, Beijing 100871, China*

Received June 6, 2022; accepted September 26, 2022; published online November 11, 2022

Although distinctive neural and physiological states are suggested to underlie the six basic emotions, basic emotions are often indistinguishable from functional magnetic resonance imaging (fMRI) voxelwise activation (VA) patterns. Here, we hypothesize that functional connectivity (FC) patterns across brain regions may contain emotion-representation information beyond VA patterns. We collected whole-brain fMRI data while human participants viewed pictures of faces expressing one of the six basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) or showing neutral expressions. We obtained FC patterns for each emotion across brain regions over the whole brain and applied multivariate pattern decoding to decode emotions in the FC pattern representation space. Our results showed that the whole-brain FC patterns successfully classified not only the six basic emotions from neutral expressions but also each basic emotion from other emotions. An emotion-representation network for each basic emotion that spanned beyond the classical brain regions for emotion processing was identified. Finally, we demonstrated that within the same brain regions, FC-based decoding consistently performed better than VA-based decoding. Taken together, our findings revealed that FC patterns contained emotional information and advocated for paying further attention to the contribution of FCs to emotion processing.

decoding, basic emotions, functional connectivity, voxelwise activation, multivariate pattern analysis

Citation: Liu, C., Wang, Y., Sun, X., Wang, Y., and Fang, F. (2023). Decoding six basic emotions from brain functional connectivity patterns. Sci China Life Sci 66, 835–847. https://doi.org/10.1007/s11427-022-2206-3

INTRODUCTION

How emotions are represented in the human brain has remained a central question in affective science (Chen et al., 2021; Fang and Hu, 2021; Giordano et al., 2021). Categorical emotion models have suggested a set of basic emotions (e.g., anger, disgust, fear, happiness, sadness, and surprise), each with discrete, dedicated neural circuitry and supporting different survival functions (Ekman, 1992; Wang et al., 2019). In support of this conjecture, neuropsychological and neuroimaging studies have revealed that different brain regions are selectively involved in processing specific basic emotions, such as the amygdala for fear (Barrett et al., 2007), the insula for disgust (Wicker et al., 2003), and the right superior temporal gyrus for happiness (Vytal and Hamann, 2010). However, previous studies failed to establish distinctive neural representations within or across brain regions that differentiate the six basic emotions (Kragel and LaBar, 2016; Murphy et al., 2003). In recent years, accumulating evidence

*Corresponding authors (Fang Fang, email: ffang@pku.edu.cn; Yingying Wang, email: ywang15@zju.edu.cn)

indicates that functional interactions across brain regions may contain more robust representations of cognitive functions than voxelwise activation (VA) in single regions (Anzellotti and Coutanche, 2018; Ito et al., 2020; Song et al., 2020; Zhang et al., 2021). Considering that emotion processing is intertwined with complicated cognitive processes including perception, motivation, and attention, it is reasonable to speculate that representations of basic emotions may be supported by large-scale functional connectivity (FC) networks in the brain (Kragel and LaBar, 2016).

Brain imaging researchers have applied multivariate pattern analysis (MVPA) to decode emotions from brain activity patterns to characterize emotion representations (Saarimäki et al., 2018; Tu et al., 2021; Wu et al., 2020). MVPA decoding models treat measured brain responses as a set of pattern vectors in a high-dimensional vector space. The multivariate patterns that most previous studies used for analysis are VA patterns within single brain regions, which has revealed that discrete distributed activity patterns in cortical regions, such as the temporal (Zhang et al., 2016) and somatomotor regions (Saarimäki et al., 2016; Saarimäki et al., 2018), as well as the subcortical regions, such as the amygdala and thalamus, underlie different basic emotions. Recent studies have further demonstrated that applying VA patterns aggregated from multiple brain regions has higher emotion decoding performance than applying VA patterns in any single region (Saarimäki et al., 2016; Saarimäki et al., 2018). However, whether the superior decoding performance results from the increased emotion representation information provided by the aggregated brain regions or is simply due to the increased number of functional magnetic resonance imaging (fMRI) voxels is unclear. Notably, neural representations in different brain regions were treated independently regardless of whether VA-based MVPA was performed on single or multiple regions of interest. In other words, the information contained in the functional interactions across regions was overlooked in traditional VA-based MVPA (Naselaris et al., 2011).

In contrast to VA-based MVPA, FC-based MVPA uses FC across brain regions as multivariate patterns under the assumption that different brain regions are interdependent during cognitive processing. Although FC-based MVPA is still at its beginning stage, a few recent studies have implicated the neural representations of emotion category information in the FCs across brain regions (Palomero-Gallagher and Amunts, 2022; Pessoa, 2018; Xu et al., 2021; Zhang et al., 2019). For example, Zhang et al. performed inter-regional FC decoding on four basic emotions and identified distinctive FC patterns for different emotions that extended beyond the conventional amygdala circuits (Zhang et al., 2019). By using FC-based MVPA, Liang et al. (Liang et al., 2018) and Saarimäki et al. (Saarimäki et al., 2022) successfully decoded six basic emotions from visual and auditory stimuli above the chance level. Furthermore, Pantazatos et al. demonstrated that FC-based decoding performed better than VA-based decoding across multiple discrete brain regions in the classification of fearful faces from neutral faces (Pantazatos et al., 2012). Thus far, most of these studies have focused on a limited set of priori-defined regions of interest and a subset of basic emotion categories. No systematic study has been conducted to explore the contribution of FCs to each basic emotion.

The goal of the present study is threefold. First, we aim to provide an MVPA pipeline for emotion classification by using FC patterns over the whole brain. Second, given the growing evidence that FCs between brain regions beyond the traditional emotion-selective and face-selective regions are involved in emotion representation, we aim to identify large-scale emotion-representation networks of individual basic emotions. Third, we test whether FCs provide additional information relative to VAs in emotion representation. To achieve these goals, this study performed MVPA over whole-brain FC patterns to identify the brain regions and networks representing different basic emotions. In a block-design experiment, participants underwent fMRI scanning while watching pictures of faces expressing the six basic emotions or showing neutral expressions. FC patterns were constructed from 112 brain nodes derived from the Harvard-Oxford atlas (Desikan et al., 2006), a whole-brain functional parcellation that has been used to explore neural networks for emotion processing (Liang et al., 2018), by using a sliding-window technique for each emotion. MVPA implemented by the Random Forest (RF) model was performed on FC patterns to classify the six basic emotions from neutral expressions and to classify each emotion from others.

RESULTS

FC-based decoding of six basic emotions

We constructed 85 (from the 20 s block design) or 61 (from the 24 s block design) FC samples for each emotion independently for the training and test sets. A binary classifier was used to decode one of the six basic emotions from neutral expressions (Figure 1A), and a seven-category classifier was utilized to decode the seven emotions (six basic emotions plus the neutral expression) (Figure 1B). As shown in Figure 1A, the binary decoding accuracies across all participants were significantly higher than the chance level (50%) for all the six basic emotions (anger, $t[11]=7.57$, $P<0.001$, Cohen's $d=2.19$; disgust, $t[11]=11.29$, $P<0.001$, Cohen's $d=3.26$; fear, $t[11]=10.41$, $P<0.001$, Cohen's $d=3.00$; happiness, $t[11]=5.58$, $P<0.001$, Cohen's $d=1.61$; sadness, $t[11]=7.31$, $P<0.001$, Cohen's $d=2.11$; surprise, $t[11]=7.80$, $P<0.001$, Cohen's $d=2.25$). The decoding accuracy of the seven-category classifier was 28.16%, which was

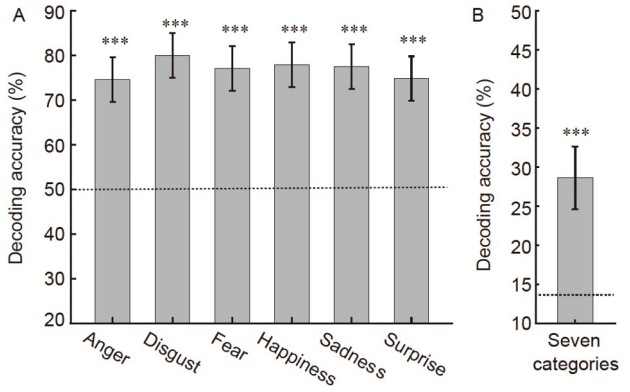

Figure 1 Accuracies of emotion decoding by using FC patterns. A, Accuracies of binary decoding classifying each of the six basic emotions from neutral expressions. B, Accuracy of classifying the seven emotions from each other. Error bars denote SEMs; ***, P<0.001 (two-tailed t-tests). Dotted lines denote the chance level.

also significantly higher (t[11]=10.00, P<0.001, Cohen's d=2.89) than the chance level (14.29%). Therefore, the whole-brain FC patterns contain information that can classify the seven emotions.

Major contributing FCs in basic emotion decoding

Next, we evaluated the contribution of the FCs to the classification of each basic emotion from neutral expressions. The emotion classifier generated a weight value for each FC during emotion classification. For evaluating the contribution of FC, we calculated the mean weight value for each FC and each emotion across all participants; this index has been used as the screening criteria for category classification (Wang et al., 2016). Then, we combined the FC weight values between brain nodes to constitute a representation network for each emotion. Figure 2 shows that the emotion-representation networks varied greatly across the six basic emotions. For example, in comparison with other basic emotions, the recognition of fearful expression (Figure 2C) recruited FCs between the limbic system (and especially the amygdala) and other brain regions; this result is in line with the critical role of the amygdala in fear perception (Adolphs, 2008; Pessoa and Adolphs, 2010). The six emotion-representation networks also showed commonalities. They commonly recruited areas for visual processing (e.g., the occipital pole), face processing (e.g., the fusiform gyrus), emotion processing (e.g., the limbic regions), and the conjunction (e.g., the supramarginal area).

Major contributing brain regions in FC-based decoding

The brain network has been demonstrated to be a small-world structure that is characterized by a combination of dense local connections and critical long-distance connections (Liu et al., 2018b; Stam and van Straaten, 2012). Therefore, only a few densely connected brain regions may be critical for representing basic emotions. We searched for brain nodes showing the densest connections with other brain nodes in emotion decoding to identify these brain regions. Specifically, we calculated the mean weight values across all participants for all FCs of each brain node and binarized them (for FC weight<0.001, we set the weight to 0; for FC weight≥0.001, we set the weight to 1). The contribution of each node to emotion classification in the FC-based decoding was quantified as the sum of the weight values of that node with all other nodes. The top 10 nodes with the greatest contribution to each basic emotion are listed in Table 1. We found that the brain nodes contributing to all six emotions were mainly in the fusiform gyrus and the occipital pole, reflecting their significance in face perception. In addition, unique brain regions were identified for the representation of individual emotions. For example, FCs from the right amygdala contributed the most to fear vs. neutral classification, reflecting that the amygdala is the central brain region for fear processing. To further display the most contributing brain regions, we applied the BrainNet view tool to project the top 10 contributing brain nodes onto the brain. As shown in Figure 3, the contributing brain nodes were mainly located in the occipital and temporal lobes.

Optimal number of brain regions for FC-based decoding

We compared the FC-based decoding accuracies obtained by using different numbers of brain nodes to identify the optimal number of brain regions for basic emotion decoding. As illustrated above, brain nodes showing the densest FCs in the whole brain were considered to have the highest contribution to emotion decoding. We ranked the 112 brain nodes in accordance with their contributions and used FC patterns from the 5, 10, 30, 50, 70, 90, and 112 brain nodes with the highest contributions to decode emotions. We found that the decoding accuracies varied significantly across different numbers of nodes for all emotions. Specifically, a 6 (number of emotions)×7 (number of nodes) repeated measures ANOVA showed a significant main effect of the number of nodes (F[6,66]=5.31, P<0.001, η^2=0.33). However, the interaction effect was not significant (F[30,330]=0.72, P=0.859, η^2=0.06). Interestingly, increasing the number of nodes did not necessarily improve the decoding performance. Instead, while the decoding accuracies increased at least marginally significantly from the 5 to 10 most contributing nodes for all the six emotions (Figure 4, anger: t[11]=2.43, P=0.033, Cohen's d=0.70; disgust: t[11]=2.37, P=0.037, Cohen's d=0.68; fear: t[11]=2.52, P=0.029, Cohen's d=0.73; happiness: t[11]=3.52, P=0.005, Cohen's d=1.02; sadness: t[11]=2.19, P=0.051, Cohen's d=0.63; surprise: t[11]=2.02, P=0.069, Cohen's d=0.58), and the accuracy decreased when additional nodes were included.

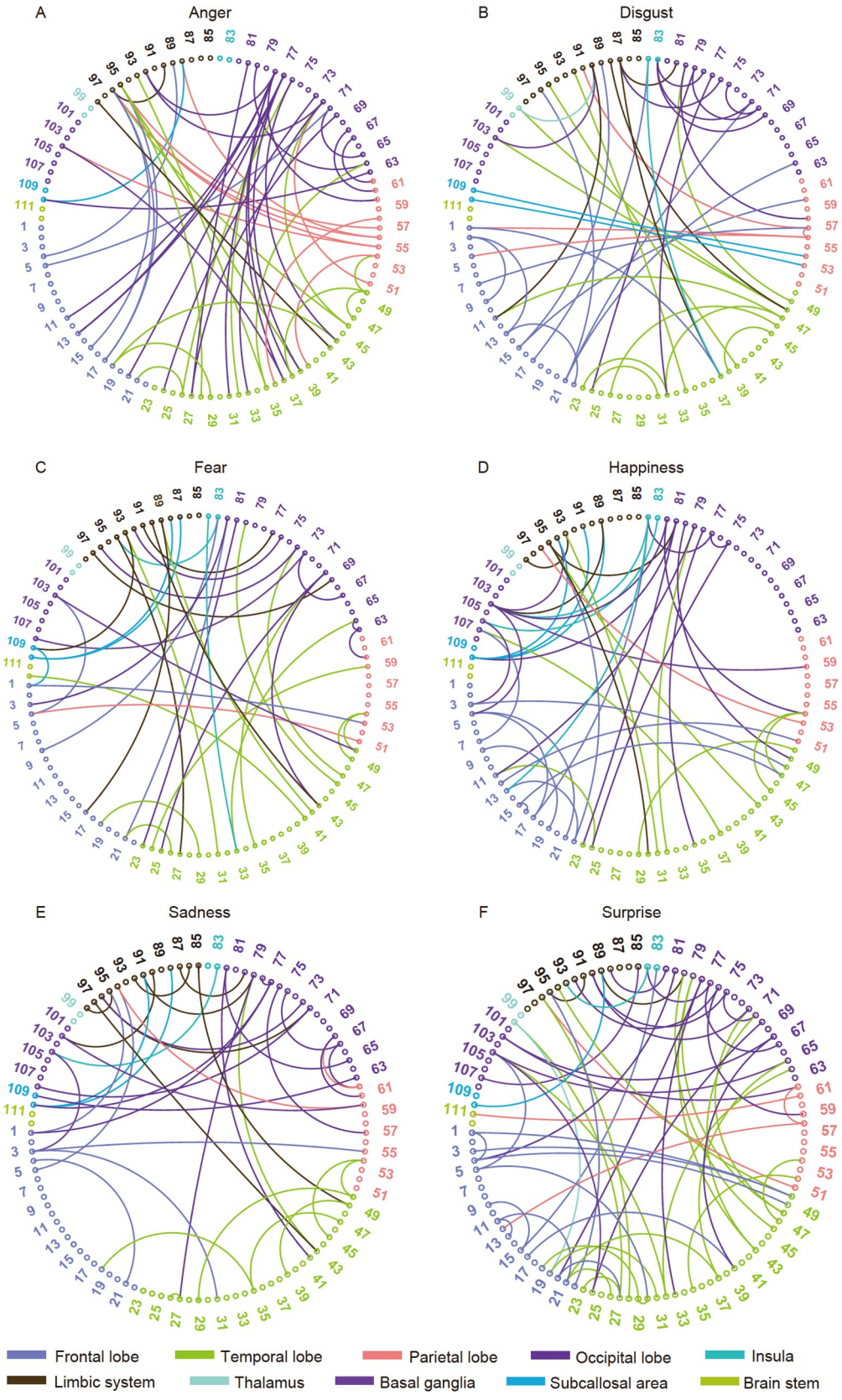

Figure 2 Emotion-representation networks of each basic emotion. FC patterns contributing to the classification of anger (A), disgust (B), fear (C), happiness (D), sadness (E), and surprise (F) from neutral expressions are shown by lines connecting different brain nodes (denoted by circles). Nodes belonging to the same brain region are shown in the same color. Each node is labeled with a number. Information for the labeling can be found in Table S1 in Supporting Information.

Table 1 Top 10 brain regions with the highest contributions to emotion decoding

Emotions	Nodes	Emotions	Nodes
Anger	Temporal occipital fusiform (R), occipital fusiform gyrus (R), heschl's gyrus (R), inferior temporal gyrus (R), supracalcarine cortex (R), putamen (R), intracalcarine cortex (R), cuneal cortex (R), occipital pole (L), occipital pole (R)	Happiness	Frontal medial cortex (L), subcallosal cortex (R), putamen (R), superior temporal gyrus (R), planum temporale (R), intracalcarine cortex (R), parahippocampal gyrus (R), occipital fusiform gyrus (R), supracalcarine cortex (R), occipital pole (R)
Disgust	Subcallosal cortex (L), occipital pole (R), inferior temporal gyrus (L), lateral occipital cortex (R), lingual gyrus (R), supramarginal gyrus (R), lateral occipital cortex (L), occipital fusiform gyrus (L), inferior frontal gyrus (L), occipital pole (L)	Sadness	Inferior frontal gyrus (R), intracalcarine cortex (R), frontal medial cortex (L), frontal medial cortex (R), heschls gyrus (L), central opercular cortex (L), supracalcarine cortex (R), occipital fusiform gyrus (R), occipital pole (L), occipital pole (R)
Fear	Amygdala (R), occipital fusiform gyrus (R), inferior temporal gyrust (L), parietal operculum cortex (L), supracalcarine cortex (R), putamen (R), superior temporal gyrus (L), occipital pole (L), cuneal cortex (R), occipital pole (R)	Surprise	Supracalcarine cortex (L), brain-stem (L), intracalcarine cortex (R), occipital fusiform gyrus (R), planum temporale (R), inferior frontal gyrus (R), superior temporal gyrus (L), occipital pole (L), supracalcarine cortex (R), occipital pole (R)

This pattern was consistently observed for the six emotions and was most evident in the happiness vs. neutral decoding (10 vs. 30, $t[11]=2.71$, $P=0.021$, Cohen's $d=0.78$; 10 vs. 50, $t[11]=2.28$, $P=0.044$, Cohen's $d=0.66$; 10 vs. 70, $t[11]=2.48$, $P=0.030$, Cohen's $d=0.72$; 10 vs. 90, $t[11]=2.41$, $P=0.041$, Cohen's $d=0.67$; 10 vs. 112, $t[11]=2.18$, $P=0.052$, Cohen's $d=0.63$). These results indicated that FC patterns with a relatively small number of nodes contain critical emotional information.

Comparison between FC-based and VA-based decoding

While previous MVPA studies using VA-based decoding often failed to decode basic emotions, our FC-based decoding achieved high decoding accuracies. However, in contrast to previous studies using VA information only within a single brain node, our combination of multiple nodes in FC patterns utilized not only FCs but also included additional VA information in the decoding model. Therefore, whether our high FC-based decoding accuracies resulted from the additional VA information from multiple brain regions or the additional information contained in FCs is unclear. We performed FC-based and VA-based MVPAs with the same combination of brain nodes, thus matching the amount of information in the two MVPA models, to address this issue. Then, we compared the decoding accuracies from FC-based decoding with those from VA-based decoding.

The FC-based decoding accuracies of the 112 nodes (Ps<0.01, Bonferroni corrected) and the VA-based decoding accuracies of the 112 combined nodes (Ps<0.01, Bonferroni corrected) were consistently significantly above the chance level (Figure 5). Moreover, the decoding accuracies obtained by using FC patterns were significantly higher than those obtained by using VA patterns (Figure 5A, anger: $t[11]=2.74$,

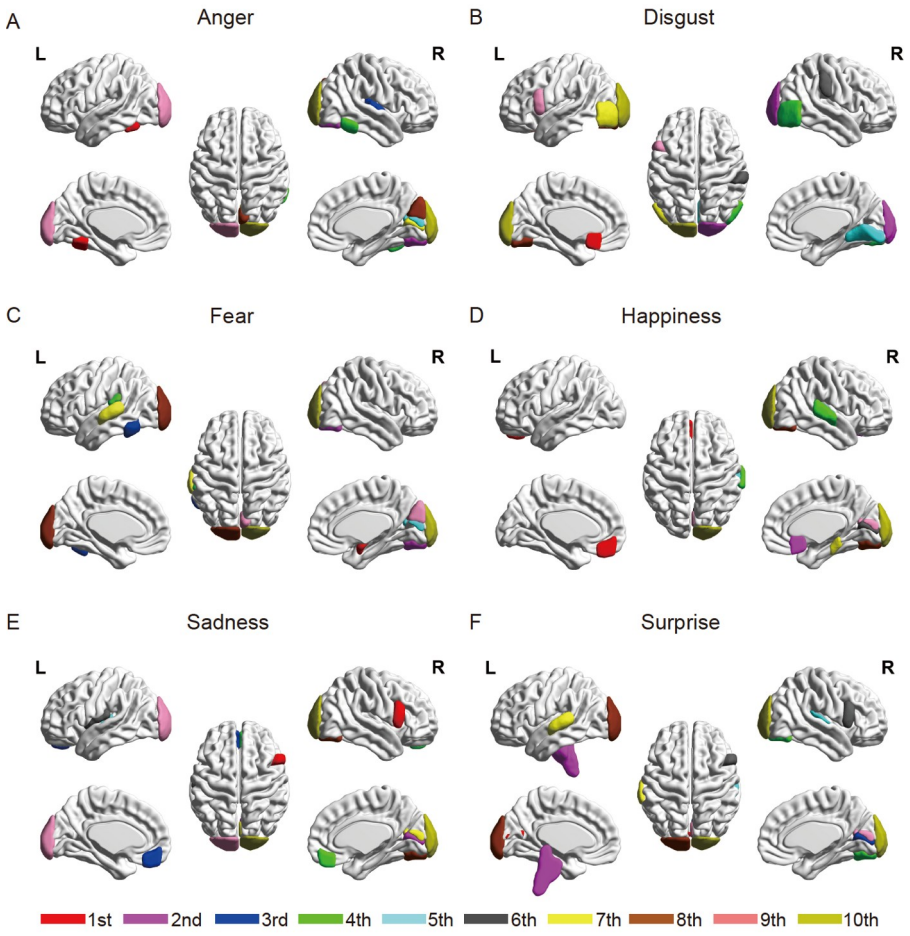

Figure 3 Distributions of the top 10 contributing brain nodes for each basic emotion. The top 10 nodes are depicted in different colors, as illustrated at the bottom of the figure.

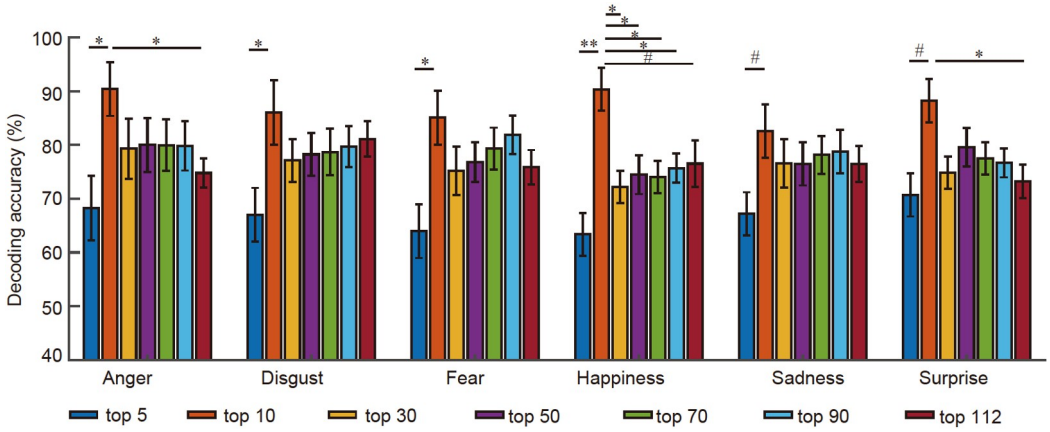

Figure 4 Effect of node number on decoding accuracy. Decoding accuracies peak at approximately ten nodes. Including additional brain regions does not increase FC-based decoding accuracies. #, P<0.07; *, P<0.05; **, P<0.01 (two-tailed t-tests). Error bars represent SEMs.

P=0.019, Cohen's d=0.79; disgust: t[11]=6.90, P<0.001, Cohen's d=1.99; fear: t[11]=6.89, P<0.001, Cohen's d=1.99; happiness: t[11]=2.95, P=0.013, Cohen's d=0.85; sadness: t[11]=3.63, P=0.004, Cohen's d=1.05; surprise: t[11]=4.51, P<0.001, Cohen's d=1.30). Next, we applied the same analysis to the top 10 nodes with the highest contributions. Replicating the whole-brain-analysis results, Figure 5B revealed that the FC-based decoding accuracies were significantly higher than the VA-based decoding accuracies (Figure 5B, anger: t[11]=5.72, P<0.001, Cohen's d=1.65;

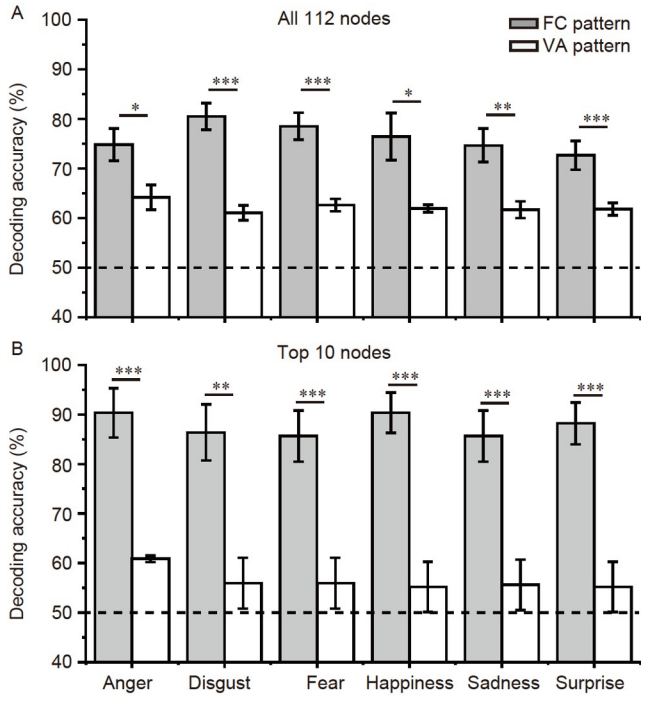

Figure 5 Comparison of FC-based decoding and VA-based decoding. A, FC patterns among all 112 nodes show higher decoding accuracies than the VA patterns from the combination of these nodes consistently for the six basic emotions. B, FC patterns of the top 10 contributing nodes consistently show higher decoding accuracies than the VA patterns of the top 10 contributing nodes for the six basic emotions. **, $P<0.01$, ***, $P<0.001$ (two-tailed *t*-tests). Error bars represent SEMs.

disgust: $t[11]=3.71$, $P=0.003$, Cohen's $d=1.07$; fear: $t[11]=6.81$, $P<0.001$, Cohen's $d=1.96$; happiness: $t[11]=5.16$, $P<0.001$, Cohen's $d=1.49$; sadness: $t[11]=6.66$, $P<0.001$, Cohen's $d=1.92$; surprise: $t[11]=5.01$, $P<0.001$, Cohen's $d=1.45$). Interestingly, only the VA-based decoding accuracy for anger (but not the other emotions) was significantly above the chance level ($t[11]=16.32$, $P<0.001$, Cohen's $d=4.71$). Overall, the FC patterns showed superior decoding performance to the VA patterns.

DISCUSSION

Our FC-based decoding method, which uses a sliding window technique and the RF model, classifies the six basic emotions from each other and classifies each basic emotion from neutral expressions with high accuracy, thus providing an effective decoding pipeline for emotion classification. Furthermore, this method identifies the emotion-representation network and the major contributing brain regions for each basic emotion. These networks not only contain classical brain regions for basic emotion representations but also reveal previously overlooked emotion-representation regions. Finally, our results add to the recently growing body of evidence that FC patterns contain more useful information for emotion decoding than VA patterns (Hutchison et al., 2013; Wong et al., 2021).

Most MVPA studies have used VA patterns of single or across multiple brain regions to classify emotions (Liang et al., 2017; Liang and Liu, 2020; Saarimäki et al., 2022). The critical finding that multiregion decoding outperforms single-region decoding has been used to attribute the contribution of functional interactions across brain regions to decoding performance (Saarimäki et al., 2022). However, this claim is not necessarily true because combining multiple brain regions simply introduces more voxelwise information than a single region. A few studies using FC-based decoding have shown the successful classification of emotions, thus providing preliminary evidence that FCs contain emotion category information (Liang et al., 2018; Palomero-Gallagher and Amunts, 2022). Here, we demonstrated the successful decoding of the six basic emotions from whole-brain FC patterns. Moreover, we directly compared the decoding performance based on the FC patterns of a selected set of brain nodes with that based on VA patterns in the same combined nodes to test whether the successful decoding from FC patterns resulted from the inclusion of additional voxels in the decoding model. We demonstrated that despite having the same amount of voxelwise information, FC-based decoding outperformed VA decoding. Therefore, FC patterns contain additional discriminant information for emotion categories (Pantazatos et al., 2012).

The high emotion classification accuracy obtained from FC patterns may benefit from several manipulations. First, we concatenated stimulus blocks and adopted a sliding window technique to construct FC patterns for each emotion. FCs computed from the concatenated same emotion blocks across different runs reflect the processing of a given emotion better than FCs from a single block (Henriksson et al., 2015; Liu et al., 2018a) and provide an increased number of time points for prediction (Cole et al., 2021). Second, we adopted the RF model, an ensemble classifier that has been suggested to boost decoding accuracy (Denisko and Hoffman, 2018). Third, we constructed FC patterns from whole-brain regions and then identified the contribution of individual brain regions based on a small network.

The emotion-representation networks contain face-selective and conventional emotion-selective brain regions. Considerable attention has been paid to the role of face-selective areas in classifying facial expression-based emotions by using VA-based decoding (Liang et al., 2017). In line with previous findings, we found that face-selective regions, including the occipital pole and the fusiform gyrus, were involved in representing all the six basic emotions (Figure 2 and Table 1). Likewise, emotion-selective regions selectively discriminated their corresponding emotions. For example, the right amygdala ranked first in the contribution to the decoding of fear, which is consistent with the key role

of the amygdala in fear processing (Adolphs, 2008). Brain regions whose FCs with the amygdala discriminated fearful from neutral faces were mainly located in the visual pathway, including the middle temporal gyrus, inferior temporal gyrus, parahippocampal gyrus, lateral occipital cortex, and occipital cortex. Fear vs. neutral discrimination also relies on the FC between the amygdala and inferior frontal gyrus, which has been shown to be involved in emotional regulation (Cha et al., 2016). We found that FCs between the putamen and the superior temporal gyrus play a key role in the decoding of happiness. Discriminating happiness from neutral expressions also relies heavily on FCs between the putamen and frontal regions, including the superior frontal gyrus, middle frontal gyrus, and frontal operculum cortex. These findings are consistent with the results of previous models and meta-analyses that separately identified these regions as the primary neural substrates for processing happy emotion (Ceravolo et al., 2021; Vytal and Hamann, 2010; Wang et al., 2014; Zhang et al., 2016).

The emotion-representation networks include additional brain regions beyond the conventional face-selective and emotion-selective regions for the representation of each basic emotion. Some of the top 10 contributing brain nodes in our results, including the supramarginal gyrus, cuneal cortex, frontal medial cortex, supracalcarine cortex, and helschl's gyrus, have rarely been claimed to contribute to emotion or face representations. These new brain regions provide evidence that different brain networks underlie the six basic emotions. However, because their contribution was derived from functional interactions with other brain regions, we should be careful about speculating the function of these brain regions *per se*. One interesting finding is that including FC patterns from additional brain nodes did not necessarily improve decoding performance. In this work, the decoding accuracies for all the six basic emotions peaked when the ten nodes with the highest contributions were used. Pantazatos et al. compared the FC-based decoding accuracy for subconscious fear vs. neutral classification across different numbers of brain regions and found a similar trend wherein classification accuracy peaked at approximately 25 brain regions (Pantazatos et al., 2012). These findings indicated that a relatively small number of brain regions contain critical information for emotion classification.

This study has some limitations. First, the decoding model only describes the data structure for different basic emotions; it does not provide a causal explanation for the representations of the basic emotions or a prediction for the representation of other emotions (Shmueli, 2010). Second, while our findings provide support for the categorical model of emotion processing, they do not exclude other frameworks, such as dimensional theories (Hamann, 2012) and mixed emotions (Hu et al., 2019), which suggest that emotions arise from the combinations of fundamental dimensions, including arousal and valence or positive and negative activation. How such emotional dimensions are represented in the brain remains unclear. Third, given that all participants in this study are Chinese, emotion perception from faces in the NimStim dataset may suffer from the influence of cultural differences. However, this influence should be small given the very similar emotion recognition performance on faces from the NimStim dataset and Chinese faces at the behavioral level (Wang et al., 2019). Fourth, despite achieving good decoding results, the effectiveness of the sliding window technique is affected by the window length and step length, and we only took into account the time-invariant information in the present model. Such an approach may limit decoding performance. Future studies may benefit from considering the dynamic information contained by FC patterns. Finally, the relatively small sample size of 12 participants may be argued to limit the reproducibility of the present findings using independent or large samples. This is unlikely. In fact, the decoding results are remarkably consistent across participants, as shown by the small SEMs across participants (all being less than 5) and the high significance level of the decoding accuracies (two-tailed *t*-test: $P<0.001$). Furthermore, the sample size of this work is comparable with that of previous studies (Szucs and Ioannidis, 2020).

CONCLUSION

By leveraging the sliding window technique and the RF model, this study provides evidence that FC patterns contain the representational information of basic emotions. Our results highlight that emotion decoding from FC patterns is superior to that from VA patterns in combined brain regions. Overall, this study suggests a potential mechanism that employs the interactions between distributed brain regions for facial expression recognition. Future studies may obtain additional information on emotion representations by studying FC patterns.

MATERIALS AND METHODS

Participants

Thirteen healthy participants (eight females; aged 18–24) were recruited from Peking University, China. One participant was excluded due to excessive head movements. All participants were right-handed, had a normal or corrected-to-normal vision, and had no past neurological or psychiatric history. They provided informed consent in accordance with the protocol approved by the human subject review committee of Peking University.

Stimuli

Stimuli were face images selected from the NimStim dataset (Tottenham et al., 2009). The faces belonged to 24 different individuals (12 males and 12 females), and each individual expressed the six basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) and a neutral expression (Wang et al., 2019). All face images were cropped with an oval region to exclude most of the hair and background, converted into grayscale, and normalized to have equivalent sizes, luminances, and RMS contrasts beforehand. They were presented on a gray background (luminance: 12.5 cd m^{-2}). The images were back-projected by using a video projector (refresh rate: 60 Hz; spatial resolution: 1024×768) onto a translucent screen placed inside the scanner bore. The participants viewed the stimuli through a mirror located above their eyes. The images subtended a visual angle of 8°×6.3°.

Experimental procedure

The experiment was completed in 2 days. Each daily session consisted of 10 (for 10 participants) or eight (for two participants) runs. A block design was used (Figure 1A left). Each run contained seven stimulus blocks, one for each emotion. Each stimulus block was immediately followed by a blank block with the same duration as the stimulus block. The block duration was 20 s for 10 participants who completed 10 runs per session and 24 s for two participants who completed eight runs per session.

In each block, images of faces expressing one of the six basic emotions or a neutral expression were presented sequentially on the screen, each for 0.8 s, followed by a 0.2 s interval. A fixation cross was presented in the screen center throughout the whole experiment. To avoid visual adaptation, the center of the face images was located at a random location within a circular region with a radius of 5° around the central fixation. The center of each face image was displayed offset by 2° relative to that of the preceding image. Participants detected repeated faces while maintaining fixation. One to three repetitions were randomly assigned to each block.

fMRI data acquisition

Whole-brain MRI data were collected by using a 3.0T Siemens Prisma MRI scanner at the Center for MRI Research at Peking University. A high-resolution simultaneous multi-slice EPI sequence was used for functional scanning (64 horizontal slices; TR=2000 ms; TE=30 ms; flip angle=90°; FOV=224×224 mm; matrix=112×112; slice thickness=2 mm; slice gap=0.2 mm; GRAPPA factor=2; multiband acceleration factor=2). A T1-weighted high-resolution 3D structural dataset was acquired for each participant in each session separately by using a 3D-MPRAGE sequence (voxel size: 1 mm×1 mm×1 mm).

fMRI data preprocessing

The preprocessing of functional images, including realignment, normalization, and smoothing, was performed by using SPM8 (Welcome Department of Cognitive Neurology, University College of London, London, UK) and the DPABI toolbox (Yan et al., 2016). The fMRI images were aligned to the first volume of the first scanning run. The images were then normalized to the Montreal Neurological Institute template and smoothed with a 4-mm full-width at half maximum Gaussian kernel.

FC-based decoding

Our decoding of emotions from fMRI data can be described by using the following formula:

$$f: fMRI\_sequence(t_1-t_2) \rightarrow k, \quad (1)$$

where t_1, t_2 are two different time points, and k is the emotion category of the faces viewed by the subject. This formula maps the FC patterns between t_1 and t_2 to an emotion category.

Sliding window

Sliding window analysis is one of the most frequently used methods for investigating the various characteristics and implications of FCs (Baczkowski et al., 2017; Jiang et al., 2022). For the acquisition of FC patterns with increased informativeness and stability (Liu et al., 2020; Zhuang et al., 2020), we concatenated all TRs corresponding to the same emotion to generate long fMRI time series (Pantazatos et al., 2012) and utilized the sliding window technique to calculate connectivity information (Figure 1B) (Baczkowski et al., 2017). The number of sliding time windows is defined by the formula:

$$W = \left\lfloor \frac{(T-w)}{\beta} \right\rfloor + 1, \quad (2)$$

where T is the total number of TRs, w is the window length, and β is the step length.

The starting time point S for each window p is defined as

$$S_p = 1 + (p-1)\beta, p = 1, \ldots, W, \quad (3)$$

and the ending time point E for each window p is defined as

$$E_p = (p-1)\beta + w, p = 1, \cdots, W. \quad (4)$$

Previous studies have demonstrated that the optimal sliding window length to capture short-lived variations and stable-state patterns from task-state fMRI data is approximately 1 min (Liu et al., 2020; Zhuang et al., 2020). Meanwhile, an offset of a single TR has been shown to be sufficiently sensitive to detect state transitions and state

durations (Shakil et al., 2016). Therefore, we set the window length of the sliding window (w) to be 36 TRs and the step length of the sliding window (β) to be one TR. As a result, W was 85 and 61 for the two block lengths.

FC pattern space

The FC patterns for each emotion were estimated on the basis of the correlations between brain nodes. The pair-wise FC patterns are computed as

$$\rho_{i,j,p,w}=\frac{\sum_{s=S_p}^{s=E_p}(y_{i,s}-\tilde{u}_{i,w})(y_{j,s}-\tilde{u}_{j,w})}{\sqrt{\left(\sum_{s=S_p}^{s=E_p}(y_{i,s}-\tilde{u}_{i,w})^2\right)\left(\sum_{s=S_p}^{s=E_p}(y_{j,s}-\tilde{u}_{j,w})^2\right)}}, \quad (5)$$

where i, j denote the node indexes (which is between 1 and 112 based on the cortical and subcortical parcellation in accordance with the Harvard-Oxford atlas), and $\tilde{u}_{i,w}$ is the mean value of the time series within the sliding window w. Thus, $\rho_{i,j,p,w}$ is the FC value between the two nodes in a time window.

The FC pattern matrix for each emotion is represented by the following formula:

$$X_t^k=(\rho_{i,j,p,w})_{i,j}, \quad (6)$$

where i, j=1,…,112. X_t^k represents the FC pattern matrix for the emotion category k at the time point t. Therefore, each FC pattern matrix is a 112×112 symmetric connectivity matrix.

The entries of the lower left triangle of each FC pattern matrix, which are the connectivity values among all 112 nodes, were linearized to construct the FC pattern sample vector. The FC pattern sample is represented by x_t^k:

$$\begin{aligned}x_t^k&=linearly\left\{lowhalf\left(X_t^k\right)\right\}\\&=linearly\left\{(\rho_{i,j})_{i,j(i>j)}\right\}\\&=(\rho_{2,1},\rho_{3,1},\ldots,\rho_{112,1},\rho_{3,2},\ldots,\rho_{112,111}),\end{aligned} \quad (7)$$

The FC pattern samples for each emotion were then pooled together to yield the representation subspace for each category:

$$X^k=\left\{x_t^k|t=1,\ldots,n\right\}, X^k\subseteq\Omega^k, \quad (8)$$

where n is the number of FC pattern samples for the given emotion category k, and Ω^k denotes their brain representation space.

Classifier

Our decoding analysis was performed on FC patterns to decode different emotions. We separately constructed training and test datasets for FC pattern samples (Figure 6D) as follows:

$$\begin{aligned}Train&=\left\{(x_l,k)|l=1,\ldots,n_1\right\}, x_l\in X^k,\\Test&=\left\{(x_q,k)|q=1,\ldots,n_2\right\}, x_q\in X^k,\end{aligned} \quad (9)$$

where x_l, x_q represent FC pattern samples, k denotes the emotion category, and n_1, n_2 represent the sample numbers. In our analysis, fMRI data from two scanning days were used as the training and test datasets.

We introduced the RF model, an ensemble classifier, to identify the FC patterns of different emotions. RF uses decision trees to perform the recursive partitioning of the feature space to yield a decision boundary and can detect stable distributed brain activation patterns (Langs et al., 2011). The objective of building decision trees is to obtain a set of decision rules that can be used to predict the category of a set of input variables. At each decision node of a decision tree, the CART algorithm was used to seek cut-off points by minimizing the Gini value under the continuous variable function. The Gini value was computed by using the following formula:

$$i(x_t)=\sum_{k\neq c}p(k|x_t)p(c|x_t), \quad (10)$$

where k, c denote the emotion categories, and $p(k|x_t)$ is the conditional probability of category k for the FC pattern sample x_t.

The training process of the decoding model was as follows: (i) The FC samples were randomly selected from the training pool to yield a small training set; (ii) an arbitrary number of q attributes was randomly selected from the high-dimensional feature attributes to train a decision tree classifier by using the CART methodology; and (iii) an RF model containing N classifiers was obtained by repeating the above process N times. The emotion category was then voted by the N classifiers. The RF model was implemented by using sklearn on Python 3.8 with the following parameters: boot_strap=true, criterion=Gine value, max_feature=sqrt, min_impurity_decrease=0.0, min_sample_leaf=1, min_samples_split=2, n_estimators (N)=40, n_jobs=1.

VA-based decoding

VA-based decoding built the machine learning model based on voxelwise activation patterns. We adopted the VA-based decoding method with the same brain nodes as those used in FC-based decoding to compare the decoding effects of the FC-based decoding method with those of the VA-based decoding method.

First, 112 brain node masks were produced for each participant in accordance with the Harvard-Oxford atlas:

$$\overline{W}=(w_1,w_2,\ldots,w_{112}), \quad (11)$$

where $\overline{W}$ represents the brain node mask set, and w_i represents the individual brain node mask.

The VA pattern in each brain node was obtained through an element-by-element multiplication operation between the preprocessed images and the corresponding brain mask:

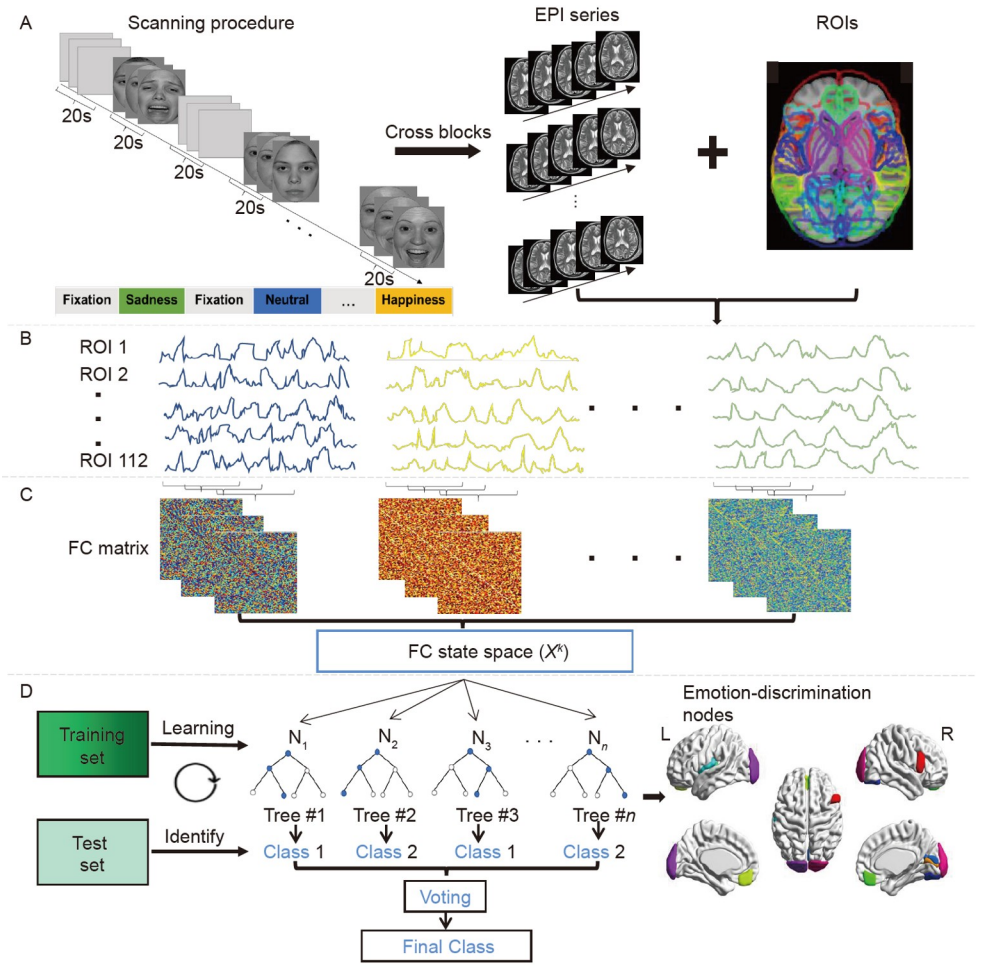

Figure 6 Flowchart of the FC decoding procedure. A, (left) Schematic of the repetition detection task. Faces images showing each of the six basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) or a neutral expression were presented sequentially in each block. Participants maintained fixation while performing the one-back task on face identity. (middle) Functional volumes were extracted for each emotion category. (right) A total of 112 brain nodes were defined in accordance with the Harvard-Oxford atlas. B, Time series of 112 nodes for each emotion was extracted. C, FC matrices were constructed by calculating the correlation coefficients across 112 nodes. The lower left triangles of the FC matrices for all emotions were pooled together to construct the FC representation spaces. D, A cross-validation procedure (left) was used to decode FC patterns with the RF model (middle) and identify the emotion-discrimination brain nodes (right) in accordance with the ranks of the contributing nodes from the classifier.

$$x_{v_i} = w_i * X, \tag{12}$$

where X represents fMRI time series data.

The VA pattern across different brain nodes (x_v) can be obtained from the following formula:

$$x_v = (x_{v_1}, x_{v_2}, \ldots, x_{v_j})_{j \leqslant 112}, \tag{13}$$

where i represents the brain node index.

These fMRI time series data were then reorganized into a new matrix ($M \times N$), where M is the number of scans and N is the number of selected voxels. Here, the selected voxels were treated as features, and the volumes were treated as samples in the MVPA model.

Second, we divided the fMRI data into two sets, namely, the training and the test sets. Similar to that in the FC-based decoding, fMRI data obtained from the two scanning days were used as the training and test datasets in this step. Thus, we had 96 or 100 samples for each emotion category in the training and test sets, respectively.

Lastly, we adopted the same RF classifier model as that used in FC-based decoding. The parameter settings of the classifier model were the same as those of the FC-based decoding model.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National Natural Science Foundation of China (31930053), the National Science and Technology Innovation 2030 Major Program (2022ZD0204802), Beijing Academy of Artificial Intelligence (BAAI), Project funded by China Postdoctoral Science Foundation (2022M710210), and the Fundamental Research Funds for the Central Universities (2021FZZX001-06). We also thank Xin-Yue Yang and Ruolin Yang for giving ideas during writing paper.*

References

Adolphs, R. (2008). Fear, faces, and the human amygdala. Curr Opin Neurobiol 18, 166–172.

Anzellotti, S., and Coutanche, M.N. (2018). Beyond functional connectivity: investigating networks of multivariate representations. Trends Cogn Sci 22, 258–269.

Baczkowski, B.M., Johnstone, T., Walter, H., Erk, S., and Veer, I.M. (2017). Sliding-window analysis tracks fluctuations in amygdala functional connectivity associated with physiological arousal and vigilance during fear conditioning. NeuroImage 153, 168–178.

Barrett, L.F., Bliss-Moreau, E., Duncan, S.L., Rauch, S.L., and Wright, C.I. (2007). The amygdala and the experience of affect. Soc Cogn Affect Neurosci 2, 73–83.

Ceravolo, L., Frühholz, S., Pierce, J., Grandjean, D., and Péron, J. (2021). Basal ganglia and cerebellum contributions to vocal emotion processing as revealed by high-resolution fMRI. Sci Rep 11, 10645.

Cha, J., DeDora, D., Nedic, S., Ide, J., Greenberg, T., Hajcak, G., and Mujica-Parodi, L.R. (2016). Clinically anxious individuals show disrupted feedback between inferior frontal gyrus and prefrontal-limbic control circuit. J Neurosci 36, 4708–4718.

Chen, J., Wang, Z., Li, Z., Peng, D., and Fang, Y. (2021). Disturbances of affective cognition in mood disorders. Sci China Life Sci 64, 938–941.

Cole, M.W., Ito, T., Cocuzza, C., and Sanchez-Romero, R. (2021). The functional relevance of task-state functional connectivity. J Neurosci 41, 2684–2702.

Denisko, D., and Hoffman, M.M. (2018). Classification and interaction in random forests. Proc Natl Acad Sci USA 115, 1690–1692.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980.

Ekman, P. (1992). An argument for basic emotions. Cognition Emotion 6, 169–200.

Fang, F., and Hu, H. (2021). Recent progress on mechanisms of human cognition and brain disorders. Sci China Life Sci 64, 843–846.

Giordano, B.L., Whiting, C., Kriegeskorte, N., Kotz, S.A., Gross, J., and Belin, P. (2021). The representational dynamics of perceived voice emotions evolve from categories to dimensions. Nat Hum Behav 5, 1203–1213.

Hamann, S. (2012). Mapping discrete and dimensional emotions onto the brain: controversies and consensus. Trends Cogn Sci 16, 458–466.

Henriksson, L., Khaligh-Razavi, S.M., Kay, K., and Kriegeskorte, N. (2015). Visual representations are dominated by intrinsic fluctuations correlated between areas. NeuroImage 114, 275–286.

Hu, X., Chen, J., Wang, F., and Zhang, D. (2019). Ten challenges for EEG-based affective computing. Brain Sci Adv 5, 1–20.

Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna, S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. NeuroImage 80, 360–378.

Ito, T., Hearne, L., Mill, R., Cocuzza, C., and Cole, M.W. (2020). Discovering the computational relevance of brain network organization. Trends Cogn Sci 24, 25–38.

Jiang, F., Jin, H., Gao, Y., Xie, X., Cummings, J., Raj, A., and Nagarajan, S. (2022). Time-varying dynamic network model for dynamic resting state functional connectivity in fMRI and MEG imaging. NeuroImage 254, 119131.

Kragel, P.A., and LaBar, K.S. (2016). Decoding the nature of emotion in the brain. Trends Cogn Sci 20, 444–455.

Langs, G., Menze, B.H., Lashkari, D., and Golland, P. (2011). Detecting stable distributed patterns of brain activation using Gini contrast. NeuroImage 56, 497–507.

Liang, Y., Liu, B., Xu, J., Zhang, G., Li, X., Wang, P., and Wang, B. (2017). Decoding facial expressions based on face-selective and motion-sensitive areas. Hum Brain Mapp 38, 3113–3125.

Liang, Y., Liu, B., Li, X., and Wang, P. (2018). Multivariate pattern classification of facial expressions based on large-scale functional connectivity. Front Hum Neurosci 12, 94.

Liang, Y., and Liu, B. (2020). Cross-subject commonality of emotion representations in dorsal motion-sensitive areas. Front Neurosci 14, 567797.

Liu, C., Song, S., Guo, X., Zhu, Z., and Zhang, J. (2018a). Image categorization from functional magnetic resonance imaging using functional connectivity. J Neurosci Methods 309, 71–80.

Liu, C., Li, Y., Song, S., and Zhang, J. (2020). Decoding disparity categories in 3-dimensional images from fMRI data using functional connectivity patterns. Cogn Neurodyn 14, 169–179.

Liu, J., Zhang, F., Liu, X., Zhuo, Z., Wei, J., Du, M., Chan, Q., Wang, X., and Wang, D. (2018b). Altered small-world, functional brain networks in patients with lower back pain. Sci China Life Sci 61, 1420–1424.

Murphy, F.C., Nimmo-Smith, I., and Lawrence, A.D. (2003). Functional neuroanatomy of emotions: a meta-analysis. Cogn Affective Behaval Neurosci 3, 207–233.

Naselaris, T., Kay, K.N., Nishimoto, S., and Gallant, J.L. (2011). Encoding and decoding in fMRI. NeuroImage 56, 400–410.

Palomero-Gallagher, N., and Amunts, K. (2022). A short review on emotion processing: a lateralized network of neuronal networks. Brain Struct Funct 227, 673–684.

Pantazatos, S.P., Talati, A., Pavlidis, P., and Hirsch, J. (2012). Decoding unattended fearful faces with whole-brain correlations: an approach to identify condition-dependent large-scale functional connectivity. PLoS Comput Biol 8, e1002441.

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. Nat Rev Neurosci 11, 773–782.

Pessoa, L. (2018). Understanding emotion with brain networks. Curr Opin Behaval Sci 19, 19–25.

Saarimäki, H., Ejtehadian, L.F., Glerean, E., Jääskeläinen, I.P., Vuilleumier, P., Sams, M., and Nummenmaa, L. (2018). Distributed affective space represents multiple emotion categories across the human brain. Soc Cogn Affective Neurosci 13, 471–482.

Saarimäki, H., Glerean, E., Smirnov, D., Mynttinen, H., Jääskeläinen, I.P., Sams, M., and Nummenmaa, L. (2022). Classification of emotion categories based on functional connectivity patterns of the human brain. NeuroImage 247, 118800.

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I.P., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M., and Nummenmaa, L. (2016). Discrete neural signatures of basic emotions. Cereb Cortex 26, 2563–2573.

Shakil, S., Lee, C.H., and Keilholz, S.D. (2016). Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states. NeuroImage 133, 111–128.

Shmueli, G. (2010). To explain or to predict? Statist Sci 25, 289–310, 222.

Song, T., Zheng, W., Song, P., and Cui, Z. (2020). EEG emotion recognition using dynamical graph convolutional neural networks. IEEE Trans Affective Comput 11, 532–541.

Stam, C.J., and van Straaten, E.C.W. (2012). The organization of physiological brain networks. Clin Neurophysiol 123, 1067–1087.

Szucs, D., and Ioannidis, J.P. (2020). Sample size evolution in neuroimaging research: An evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. NeuroImage 221, 117164.

Tottenham, N., Tanaka, J.W., Leon, A.C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D.J., Westerlund, A., Casey, B.J., and Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry Res 168, 242–249.

Tu, Y., Cao, J., Bi, Y., and Hu, L. (2021). Magnetic resonance imaging for chronic pain: diagnosis, manipulation, and biomarkers. Sci China Life Sci 64, 879–896.

Vytal, K., and Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. J

Cogn Neurosci 22, 2864–2885.
Wang, S., Tudusciuc, O., Mamelak, A.N., Ross, I.B., Adolphs, R., and Rutishauser, U. (2014). Neurons in the human amygdala selective for perceived emotion. Proc Natl Acad Sci USA 111, E3110–E3119.
Wang, X., Fang, Y., Cui, Z., Xu, Y., He, Y., Guo, Q., and Bi, Y. (2016). Representing object categories by connections: evidence from a mutivariate connectivity pattern classification approach. Hum Brain Mapp 37, 3685–3697.
Wang, Y., Zhu, Z., Chen, B., and Fang, F. (2019). Perceptual learning and recognition confusion reveal the underlying relationships among the six basic emotions. Cognition Emotion 33, 754–767.
Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in my insula. Neuron 40, 655–664.
Wong, W., Cabral, J., Rane, R., Ly, R., Kringelbach, M.L., and Feusner, J. D. (2021). Effects of visual attention modulation on dynamic functional connectivity during own-face viewing in body dysmorphic disorder. Neuropsychopharmacology 46, 2030–2038.
Wu, D., Li, X., and Jiang, T. (2020). Reconstruction of behavior-relevant individual brain activity: an individualized fMRI study. Sci China Life Sci 63, 410–418.
Xu, P., Peng, S., Luo, Y.J., and Gong, G. (2021). Facial expression recognition: a meta-analytic review of theoretical models and neuroimaging evidence. Neurosci BioBehaval Rev 127, 820–836.
Yan, C.G., Wang, X.D., Zuo, X.N., and Zang, Y.F. (2016). DPABI: data processing & analysis for (resting-state) brain imaging. Neuroinformatics 14, 339–351.
Zhang, G., Yu, M., Liu, Y.J., Zhao, G., Zhang, D., and Zheng, W. (2021). SparseDGCNN: recognizing emotion from multichannel EEG signals. IEEE Trans Affective Comput, doi: 10.1109/TAFFC.2021.3051332.
Zhang, H., Japee, S., Nolan, R., Chu, C., Liu, N., and Ungerleider, L.G. (2016). Face-selective regions differ in their ability to classify facial expressions. NeuroImage 130, 77–90.
Zhang, Y., Padmanabhan, A., Gross, J.J., and Menon, V. (2019). Development of human emotion circuits investigated using a big-data analytic approach: stability, reliability, and robustness. J Neurosci 39, 7155–7172.
Zhuang, X., Yang, Z., Mishra, V., Sreenivasan, K., Bernick, C., and Cordes, D. (2020). Single-scale time-dependent window-sizes in sliding-window dynamic functional connectivity analysis: A validation study. NeuroImage 220, 117111.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-022-2206-3. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.