
Gyri vs. Sulci: Disentangling Brain Core-Periphery Functional Networks via Twin-Transformer

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 The human cerebral cortex is highly convoluted into convex gyri and concave
2 sulci. It has been demonstrated that gyri and sulci are significantly different
3 in their anatomy, connectivity and function: besides exhibiting opposite shape
4 patterns, long-distance axonal fibers connected to gyri are significantly denser
5 than those connected to sulci, and neural signals on gyri are more complex in the
6 low-frequency band while sulci have more complex patterns in the high-frequency
7 band. Although accumulating evidence shows significant differences between
8 gyri and sulci, their primary roles in brain function have not been elucidated
9 yet. To solve this fundamental problem, we design a novel Twin-Transformer
10 framework to explore and unveil the unique functional roles of gyri and sulci as
11 well as their relationship and interaction in the whole brain function. Our Twin-
12 Transformer framework adopts two identical and connected (twin) Transformers
13 to model and disentangle spatial-temporal patterns of gyri and sulci: one focuses
14 on the information of gyri and the other is on sulci. The Gyro-Sulcal interactions,
15 along with the tremendous but widely existing variability across individuals, are
16 characterized and represented via a novel Gyro-Sulcal Commonality-Variability
17 Disentangled Loss (GS-CV Loss). We validated our Twin-Transformer on one of
18 the largest brain imaging datasets (HCP task-fMRI gray-ordinate dataset), for the
19 first time, to elucidate the different roles of gyri and sulci in brain function. Our
20 results suggest that gyri and sulci could work together in a core-periphery network
21 manner, that is, gyri could serve as core networks for information gathering and
22 distributing in a global manner, while sulci could serve as periphery networks
23 for specific local information processing. These findings have shed new light
24 on our fundamental understanding of the brain's basic structural and functional
25 mechanisms.

26

1 Introduction

27 The human cerebral cortex (top of Fig. 1-a) is highly convoluted into convex gyri and concave sulci
28 (Fig. 1-b). Gyri and sulci serve as the basic building blocks to make up complex cortical folding
29 patterns, and are fundamental to realize the brain's basic structural and functional mechanisms.
30 Numerous efforts have been devoted to understanding the function-anatomy patterns of gyri and
31 sulci from various perspectives, including genetics[32], cell biology[13], and neuroimaging[25]. It
32 has been demonstrated consistently that gyri and sulci are significantly different in their anatomy,
33 connectivity and function. Several studies[11, 17, 23, 31] found that the formation of gyri/sulci may
34 be closely related to the micro-structure of white matters. For example, diffusion tensor imaging (DTI)
35 derived long-distance axonal fibers connected to gyri are significantly denser than those connected to
36 sulci (bottom of Fig. 1-a). That is, the long-distance fiber terminations dominantly concentrate on
37 gyri rather than sulci, and interestingly, this phenomenon is evolutionarily preserved across different

38 primate species. Meanwhile, using functional magnetic resonance imaging (fMRI), a few functional
39 measurements that can directly reflect brain functional activities on gyri and sulci have been explored,
40 such as functional BOLD signals [25], correlation-based connectivity/interaction [8], and spatial
41 distribution of functional networks [29, 30]. Despite accumulating functional differences found
42 between gyri and sulci, their basic roles as well as their relationship and interaction in the whole
43 brain function have not been explored or elucidated yet.

44 To answer this fundamental question in brain science, we proposed a novel Twin-Transformer
45 framework (Fig. 1-c) to explore and unveil the unique functional roles of gyri and sulci. Unlike
46 traditional factorization-based approaches that assume linearity and independence, the Transformer
47 attention mechanism is an ideal backbone to characterize, represent and reveal the complex and
48 deeply buried patterns in the observed brain functional data. Our whole framework is illustrated in
49 Fig. 2. Our Twin-Transformer framework adopts two identical and connected (twin) Transformers to
50 model and disentangle spatial-temporal patterns of gyri and sulci: one focuses on the information
51 of gyri and the other focuses on sulci. To model the complex 4D (spatial-temporal) fMRI data,
52 within each transformer, we designed a spatial module and a temporal module to disentangle and
53 extract the patterns in both spatial and temporal domains from the original fMRI signals. The two
54 Transformers are connected and interact via a group of shared weights and constraints between the
55 two spatial/temporal modules. In addition, to effectively capture the Gyro-Sulcal interactions, as
56 well as the tremendous and widely existing variability across individual brains, a novel Gyro-Sulcal
57 Commonality-Variability Disentangled Loss (GS-CV Loss) is proposed to guide the training process.
58 After the model is well-trained, the functional brain networks (FBNs) and the corresponding temporal
59 activations that are specific to gyri and sulci can be recovered by the corresponding transformers. We
60 validated our Twin-Transformer on the one of the largest brain imaging datasets (HCP task-fMRI
61 gray-ordinate dataset), for the first time, to elucidate the different roles of gyri and sulci in brain
62 function. Our results suggest that gyri and sulci could work together in a core-periphery network
63 manner (Fig. 1-d), that is, gyri could serve as core networks for information gathering and distributing
64 in a global manner, while sulci could serve as periphery networks for specific local information
65 processing. These findings have shed new light on our fundamental understanding of the brain's basic
66 structural and functional mechanisms. The contributions of this paper are summarized as follows:

- 67 • We introduced a novel Twin-Transformer to represent and unveil the fundamental functional
68 roles of the two basic cortical folding patterns: gyri and sulci.
- 69 • We discovered unique functional role patterns that are specifically located on gyri (global)
70 and sulci (local).
- 71 • We found that gyri and sulci may work together in a Core-Periphery network manner: gyri
72 serve as core networks for information gathering and distributing, while the sulci serve as
73 periphery networks for specific local information processing.

74 2 Related Works

75 2.1 Gyri and Sulci

76 Gyri and sulci are the standard morphological and anatomical nomenclature of cerebral cortex
77 and are usually defined in anatomical domains [21]. Neuroscientific studies have demonstrated
78 that gyri and sulci may emerge from a complex cortical folding process, which is closely related
79 to neurodevelopment[23], cytoarchitecture[12], and cognitive functioning[19]. Moreover, specific
80 gyral-sulcal patterns have been widely reported to be closely relevant to brain neuronal processes
81 [7, 22], functional activity [35], and human behaviors[37]. Therefore, gyral-sulcal patterns play
82 important roles in brain anatomy, function, and cognition. Unveiling their fundamental roles as well
83 as their relationship and interaction in the whole brain function is of fundamental importance to
84 understand the underlying structural and functional brain mechanisms. In this paper, for the first time,
85 we proposed a novel Twin-Transformer framework to elucidate the different roles of gyri and sulci in
86 brain function.

87 2.2 Transformer

88 Since it was first proposed in 2017 [36], with its strong representation capacity, transformer and
89 its variants, such as BERT[9] and Generative Pre-trained Transformer (GPT)[2], have achieved

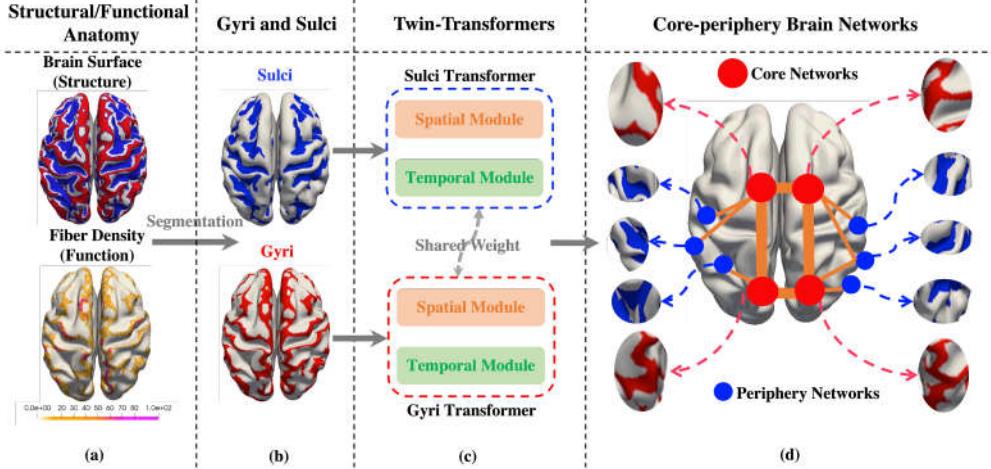


Figure 1: Core-periphery brain networks in gyri and sulci. (a) is the brain structural and functional anatomy of gyri and sulci. (b) is the segmentation of gyri and sulci. (c) is the proposed Twin-Transformer, where one is for gyri and the other is for sulci. (d) is the core-periphery brain networks derived from the gyri and sulci, where gyri is the core network, and sulci is the periphery network.

breakthroughs in the NLP domain. Inspired by the tremendous success of transformer architectures in NLP, vision transformer (ViT)[10] has been proposed by introducing transformer architecture into image representation learning and utilized to address a variety of vision tasks, such as image classification[5], object detection[3], semantic segmentation[38], image processing[4], and video understanding[39]. Thanks to its exceptional performance, transformer-based vision models have become a potential alternative to CNN in image processing domain. To leverage the brilliant spatial and temporal representation capacity of ViT in handling image/video data, we proposed a novel Twin-Transformer framework to capture the complex gyral-sulcal spatial-temporal patterns from brain function data.

2.3 Core-periphery Network

Core-periphery[33] structures are widely existing in transportation systems[34], social networks[1], financial networks[16], and brain networks[15]. The study[15] on brain complex network reported the core/periphery networks in region of interest (ROI) with fMRI, MEG and DWI data. Another study [14] demonstrated the core-periphery network universally exists in human functional brain networks and unified the core-periphery with the modular organization. However, existing studies are limited in simply reporting core-periphery structure may exist in brain newtork, the factor behind this biological phenomena is unclear. In this work, using our novel Twin-Transformer model we are able to unveil that gyri and sulci, as the two basic anatomical folding patterns, serve as the core network and periphery network, respectively.

3 Methods

3.1 Gyri and Sulci Data Preparation

In our experiments, we used high-quality task-based fMRI (tfMRI) data of 540 subjects from the Human Connectome Project (HCP), that is, 3 Tesla motor and working memory (WM) task gray-coordinate dataset[26][27]. The publicly available preprocessed tfMRI data went through the minimal preprocessing pipelines that are especially designed for high spatial and temporal resolution of HCP datasets [28]. The preprocessed tfMRI imaging data is a kind of 4D imaging data, which consists of a time-series of 3D images of the brain. For motor task-fMRI, each voxel contains a series of brain signals of length 284. We reorganize the signals in each voxel into a 2D matrix. In this way, a 4D tfMRI imaging can be represented by a 2D matrix, where rows represent the tfMRI time series and

119 columns represent the brain voxels (Fig. 2-a). We normalized the brain signals to zero mean and unit
 120 variance. Since each subject of the preprocessed data has 59,412 voxels in standard grayordinate
 121 space, the column dimension of the 2D matrix is 59,412. To facilitate patch partition, we expanded the
 122 space dimension from 59,415 to 60,000 by adding zero vectors along the spatial dimension. Finally,
 123 a set of 2D brain signal matrices of all the subjects with dimensions of 284×60,000 are generated.
 124 Then we map the gyri and sulci masks onto the 2D brain signal matrix of each subject, and both gyri
 125 and sulci signal matrices of the 284×60,000 are generated correspondingly.

126 3.2 Twin-Transformers

127 To reveal the common and variable patterns contained in the gyri and sulci, a novel Twin-Transformer
 128 framework is proposed, including a gyri transformer and a sulci transformer. The architecture of the
 129 Twin-Transformer is illustrated in Fig. 2. There is a spatial and temporal self-attention module in the
 130 gyri transformer for disentangling spatial and temporal patterns of gyri as shown in Fig. 2-c. The
 131 structure of the sulci transformer is the same as the gyri transformer. For each input signal matrix,
 132 spatial patches are generated by shifting window along the space dimension, as illustrated by the
 133 orange arrow in Fig. 2-a, while temporal patches are generated by shifting window along the time
 134 dimension, as shown in the green arrow in Fig. 2-a. Gyri transformer generates spatial and temporal
 135 patterns of brain networks on gyri, while sulci transformer generates spatial and temporal patterns of
 136 those on sulci. By constraining the spatial and temporal patterns between gyri and sulci, commonality
 137 and variability between gyri and sulci can be discovered.

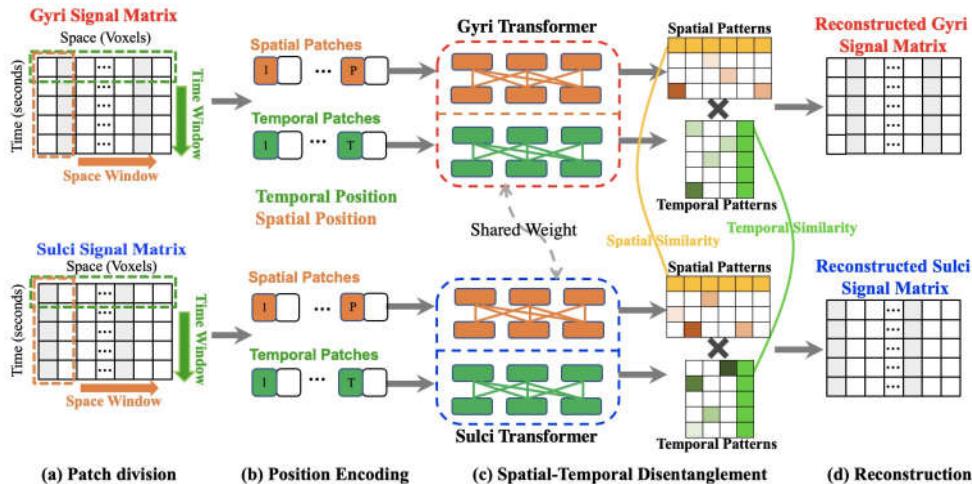


Figure 2: Illustration of the proposed Twin-Transformer framework. (a) shows the patch division of the gyri and sulci signal matrices. (b) is the position encoding for the spatial and temporal patches. (c) shows the details of the Twin-Transformer. The gyri transformer shares weights with sulci transformer, and each transformer includes a spatial and temporal self-attention module for processing spatial patches and temporal patches. (d) is the reconstruction of the gyri and sulci signal matrices from disentangled spatial and temporal patterns.

138 Specifically, within gyri or sulci transformer, the spatial self-attention module is designed to learn
 139 the latent representations of spatial features, and it focuses on the space dimension and takes non-
 140 overlapping spatial patches as tokens to build attention across the spatial variant patches and generate
 141 spatial patterns. It divides the input signal matrix into P non-overlapping patches by shifting the
 142 sliding window (orange dotted box following orange arrow) from left to right along the space
 143 dimension. The size of the sliding window can be adjusted according to the size of the input data.
 144 Each spatial patch contains complete temporal information of the focal brain region. The P patches
 145 correspond to P components of brain networks as predefined. Patches are used as tokens, and each
 146 token is first fed into a linear projection layer to obtain the latent representation $z_i \in \mathbb{R}^{1 \times D_1}$ and then
 147 the learnable spatial positional embedding, $E_i^s \in \mathbb{R}^{1 \times D_1}$ is added to the representations of each input
 148 token. The spatial transformer encoder can be formulated as:

$$Spa(Z) = MLP(MSA(LN(z_1^S || z_2^S || z_3^S || \dots || z_P^S))) \quad (1)$$

149 where MSA() is the multi-head self-attention, MLP() represents multilayer perceptron, and LN() is
 150 layernorm. $z_i^s = (z_i + E_i^S)$, $i = 1, 2, \dots, P$ and \parallel denotes the stack operation. $Spa(Z) \in P \times N$ is
 151 the output of the spatial Transformer, where P represents the number of brain networks and N is the
 152 number of voxels in the brain. $Spa(Z)$ models the activated voxels within each brain network.

153 The temporal transformer is designed to learn the latent representations of temporal patterns of
 154 brain networks. The temporal self-attention module focuses on the temporal dimension and the
 155 non-overlapping temporal patches are used as tokens. Correspondingly, the temporal Transformer
 156 builds attention across the temporal variant patches and generates temporal features. Similar to the
 157 spatial transformer, by shifting the sliding window (green dotted box following green arrow) from
 158 top to bottom along the time dimension, T non-overlapping temporal patches are generated. The size
 159 of the sliding window equals 1, hence the number of patches equals the length of the brain signals.
 160 Each temporal patch contains information of all the voxels. After input embedding and positional
 161 embedding, each patch is represented by $z_i^t = (z_i + E_i^t)$, $i = 1, 2, \dots, T$. The temporal self-attention
 162 module can be formulated as:

$$Tem(Z) = MLP(MSA(LN(z_1^t \parallel z_2^t \parallel z_3^t \parallel \dots \parallel z_P^t))) \quad (2)$$

163 The outputs $Tem(Z)$ of the temporal self-attention module have a dimension of $Tem(Z) \in T \times P$,
 164 where T represents the time points of the fMRI signals. $Tem(Z)$ represents the signal pattern of each
 165 brain network. Taking $Spa(Z)$ and $Tem(Z)$ together, we can obtain both the spatial and temporal
 166 patterns of each pair of gyri and sulci.

167 3.3 Gyri and Sulci Commonality-Variability Disentangled Loss

168 To simultaneously capture common and variable patterns in the gyri and sulci, a new gyri-sulci
 169 commonality-variability disentangled loss (GS-CV Loss) is proposed. There are three components in
 170 GS-CV Loss. The first one is the signal matrix reconstruction loss. The whole framework is trained in
 171 a self-supervised manner to reconstruct the input signal matrix from the learned spatial and temporal
 172 patterns of gyri and sulci. This is crucial to ensure the learned spatial and temporal features can
 173 capture the complete spatial and temporal information of the input data. The reconstruction loss can
 174 be formulated as:

$$L_{reco} = \sum \|X - Spa(Z) \cdot Tem(Z)\|_{L1} \quad (3)$$

175 where X is the input signal matrix, and we use L1-norm to constrain the reconstruction of the input
 176 gyri and sulci pair. The second component is the commonality constrain loss of spatial patterns
 177 between gyri and sulci, which aims to find the common spatial patterns between gyri and sulci. For
 178 this purpose, the learned spatial feature matrix is divided into common part (the first p rows) and
 179 variable part (the remaining rows). The common and variable patterns can be learned by minimizing
 180 the difference between common parts of gyri and sulci and leaving the variable parts to learn freely.
 181 This can be formulated as:

$$L_{comm_spa} = \sum Corr(\|Spa(Z_1)[-p :, *] - Spa(Z_2)[-p :, *]\|) \quad (4)$$

182 where $[0 : p, *]$ represents the first p rows in $Spa(Z_i)$, and $*$ means for each row, all the elements
 183 in the columns are included, and vice versa. Since the scale of the brain signals in gyri and sulci is
 184 different, we adopt the Pearson correlation coefficient to constrain the similarity of common spatial
 185 patterns between gyri and sulci to be maximized. Similarly, the commonality constraint on temporal
 186 features, which is the third component in GS-CV Loss, is formulated as:

$$L_{comm_tem} = \sum Corr(\|Tem(Z_1)[*, 0 : p] - Tem(Z_2)[*, 0 : p]\|) \quad (5)$$

187 In order to make spatial patterns distinct and limit the scale of temporal pattern from being arbitrarily
 188 large, we add a normalization on temporal features, which is formulated as:

$$L_{tem_norm} = \max(0, \frac{1}{P} (\sum_{i=1}^P \|Tem(Z_i[*, i])\|_2) - 1) \quad (6)$$

189 Combining the four parts, the GS-CV Loss can be formulated as:

$$GS - CV_Loss = \alpha L_{reco} + \beta L_{comm_spa} + \gamma L_{comm_tem} + \delta L_{tem_norm} \quad (7)$$

190 where the regularization parameters α, β, γ , and δ controls the balance of different factors on the
 191 overall loss function.

192 **4 Results**

193 We applied our method to one of the largest brain image dataset - HCP tfMRI data (we used both
 194 motor and working memory tasks in this work). Using the fMRI signals from gyri and sulci for each
 195 subject, as a paired input for Twin-Transformer, we generated the gyri/sulci related patterns: the
 196 output of each transformer includes 100 well-trained spatial components that can be interpreted as
 197 100 FBNs that are specific to gyri and/or sulci. The corresponding 100 temporal components can
 198 be treated as the representative signals of each FBN in the embedding space. We first illustrate the
 199 global/local patterns of gyri/sulci using both individual and group-wise results, revealing that gyri and
 200 sulci may work together in a Core-Periphery network manner. To examine the core-periphery concept
 201 in temporal domain, we further analyze the task involved rate (TIR) of the temporal components. We
 202 found that gyri have much higher TIR than sulci, which indicates that gyri participate more in tasks
 203 than sulci do. In addition, gyri dominant FBNs show clearly global distribution patterns, while sulci
 204 dominant FBNs display an opposite local mode. All of these results taken together suggest that gyri
 205 serve as core networks for information gathering and distributing, while the sulci serve as periphery
 206 networks for specific local information processing. We also tested our proposed methods on another
 207 tfMRI dataset of working memory, and the conclusions are reproducible and consistent, and these
 208 results can be found in supplementary material.

209 **4.1 Core-Periphery Network**

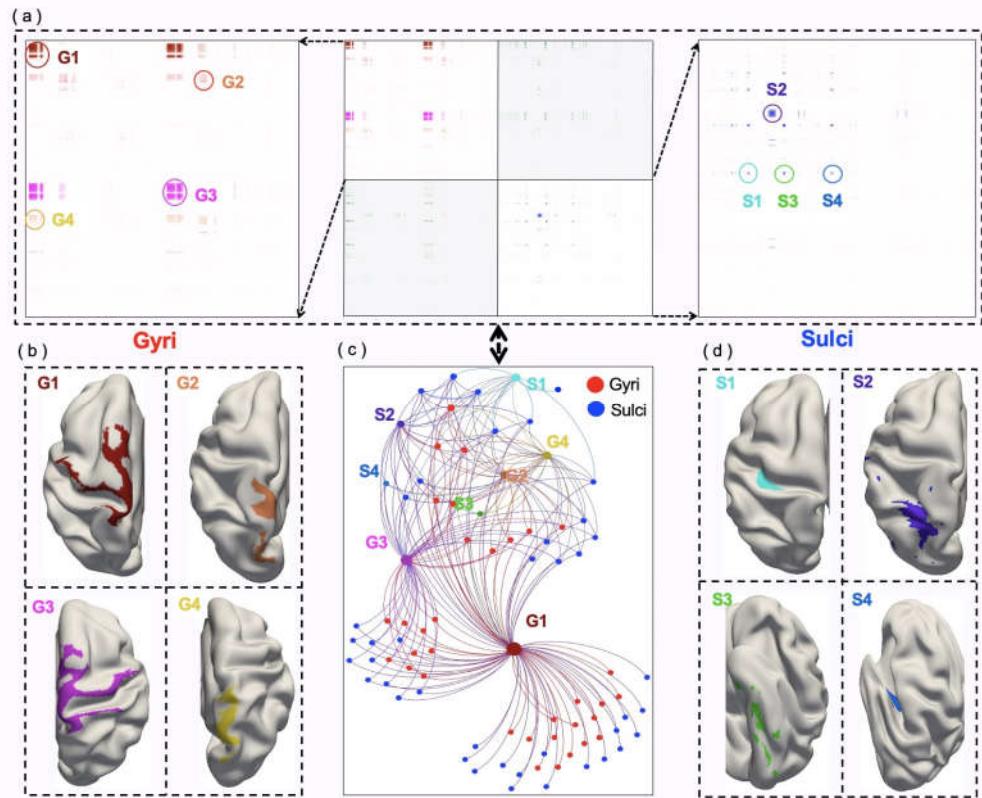


Figure 3: Core-Periphery Relationship Between Gyri and Sulci. (a): The activated voxels within gyri and sulci in common spatial brain networks. For better visualization, we enlarge the gyri and sulci parts into the left one and the right one. The major clusters of activated voxels in gyri are marked as G1-G4, whereas the major clusters in sulci noted as S1-S4. (b): The brain regions that correspond to the major activated brain voxels in gyri. The notations and colors are consistent with the gyri part in (a). (d): The brain regions that correspond to the major activated brain voxels in sulci. The notations and colors are consistent with the sulci part in (a). (c): Connected graph of the entire relationship matrix in (a). The red points are gyri, and the blue points are sulci.

Table 1: The Independent Probability of Gyri Sulci Network

IP	WM			MOTOR		
	0.10	0.15	0.20	0.10	0.15	0.20
P_{GG}	0.35 ± 0.02	0.30 ± 0.02	0.07 ± 0.05	0.42 ± 0.06	0.12 ± 0.02	0.02 ± 0.01
P_{GS}	0.20 ± 0.02	0.16 ± 0.02	0.05 ± 0.05	0.37 ± 0.05	0.08 ± 0.02	0.01 ± 0.01
P_{SS}	0.12 ± 0.02	0.09 ± 0.02	0.04 ± 0.04	0.33 ± 0.05	0.06 ± 0.02	0.01 ± 0.01

Table 2: Gyri and Sulci Ratio Under Different Experimental Settings

Components	Comm. Spatial		Comm. Temporal		Gyri-Sulci Specific	
	Gyri Ratio	Sulci Ratio	Gyri Ratio	Sulci Ratio	Gyri Ratio	Sulci Ratio
50	52.6 ± 0.08	47.4 ± 0.08	50.1 ± 0.09	49.9 ± 0.09	50.5 ± 0.03	49.5 ± 0.03
100	53.8 ± 0.08	46.2 ± 0.08	57.5 ± 0.07	42.5 ± 0.07	51.8 ± 0.03	48.2 ± 0.03
150	54.6 ± 0.05	45.4 ± 0.05	53.2 ± 0.04	46.8 ± 0.04	51.6 ± 0.04	48.4 ± 0.04
200	54.5 ± 0.01	45.5 ± 0.01	56.7 ± 0.03	43.3 ± 0.03	55.7 ± 0.04	44.3 ± 0.04

210 We can identify the activated brain voxels whose weights are consistently above a pre-defined
 211 threshold across all gyri- or sulci- derived spatial components. By connecting all the activated brain
 212 voxels, we construct a relationship matrix of gyri and sulci. Fig. 3 shows an example of one randomly
 213 selected subject (more individual cases and group-wise results have been included in supplementary
 214 material). There are 17,232 voxels for gyri and 18,327 voxels for sulci in this subject’s gray-ordinate
 215 surface, so the dimension of the obtained relationship matrix is 35559×35559 , and 17232×17232
 216 and 18327×18327 for gyri part and sulci part, respectively. The middle in Fig. 3-a demonstrates the
 217 entire relationship matrix, the sub-figures on left and right highlight the connections within gyri and
 218 sulci voxels, which are located in the top-left and bottom-right of the relationship matrix. In general,
 219 the relation matrix is sparse, which means only a few regions (voxels) are involved in a specific task
 220 at the same time, and this result is consistent with previous literature reports[20, 24, 26, 29]. The
 221 most interesting finding using our Twin-Transformer is that the activated brain voxels in gyri-gyri
 222 section (left in Fig. 3-a) incline to form larger and connected blocks or clusters, as highlighted with
 223 four circles (G1-G4), while the activated brain voxels in sulci-sulci section (right in Fig. 3-a) tend
 224 to assemble as much smaller and scattered patterns (S1-S4). It worth noting that if the voxels are
 225 close in relationship matrix, they also tend to be neighbors on cortical surface. Therefore, after
 226 mapping the blocks of G1-G4 to cortex, we can see large continuous gyri regions on the brain surface
 227 (Fig. 3-b) forming gyri-based FBNs. However, the activated regions of sulci (sulci-based FBNs) are
 228 relatively small and separated (Fig. 3-d). To further examine the relationship between gyri-based
 229 and sulci-based FBNs, we visualize the gyri-sulci section which is located in the bottom left of the
 230 relationship matrix, as a connected graph shown in Fig. 3-c. We labeled the nodes in the graph with
 231 previously identified G1-G4 and S1-S4, and build their connections according to the relationship
 232 matrix. We can clearly see that all the gyri-based FBNs serve as the hub nodes, and they together
 233 compose the Core Network. Meanwhile, the sulci-based FBNs serve the supporting nodes, forming
 234 the Periphery Network. That is, the Core Networks includes gyri-based FBNs and they connect each
 235 other directly; the Periphery Networks consist of sulci-based FBNs and their communications in the
 236 entire brain network rely on the Core Network.

237 To further prove the concept of the Core-Periphery Network of gyri and sulci, we compute the
 238 independent probability (IP) P_{GG} , P_{SS} and P_{GS} for sub-matrices A_{GG} , A_{SS} , and A_{GS} of the entire
 239 relationship matrix, which represents the interactions within gyri vertices (Core Network), sulci
 240 vertices (Periphery Network) and between gyri, and sulci vertices (between Core and Periphery
 241 Networks). Independent probability[6] is defined as the probability that there is an edge between any
 242 pairs of nodes in a given matrix, and it is an important measurement to indicate if the matrix or graph
 243 is organized as Core-Periphery pattern[18, 33]. We set three different thresholds for edge activation
 244 to calculate the IP, and the average results of 500 subjects are shown in Table 1. The results show
 245 that $P_{GG} > P_{GS} > P_{SS}$, which confirms that our derived gyri/sulci networks have the core-periphery
 246 structure.

247 **4.2 Task Involved Rates in Gyri and Sulci**

248 Besides the spatial patterns of the gyri and sulci, we examine temporal patterns of gyri and sulci in this
 249 section. We calculated the Pearson correlation coefficient (PCC) between the temporal patterns and
 250 five task stimuli in motor task: left hand, right hand, left foot, right foot, and tongue. We empirically
 251 set the threshold for PCC to consider the specific temporal pattern correlated with task stimulus. We
 252 define the task-involved rates (TIR) as the number of task stimuli that the temporal patterns involved
 253 divided by the number of all stimuli. We calculated the TIR under different experimental settings, 50,
 254 100, 150, and 200 components, and under the different thresholds for PCC. The whole TIR consists
 255 of three parts, which are common spatial TIR, common temporal TIR, and gyri-sulci specific TIR.
 256 The results are shown in Fig. 4. We can see that the TIR in gyri are all higher than that in sulci,
 257 except in common temporal patterns, since the common temporal patterns are trained to be similar
 258 under the temporal similarity loss. It has been widely recognized that a single brain task may need to
 259 recruit multiple brain regions or FBNs to work together. Our results show gyri have been involved
 260 more frequently, and in more tasks than sulci, which further indicates that gyri play a key role (Core
 261 Network) in brain activities, whereas sulci play a supportive role (Periphery Network).

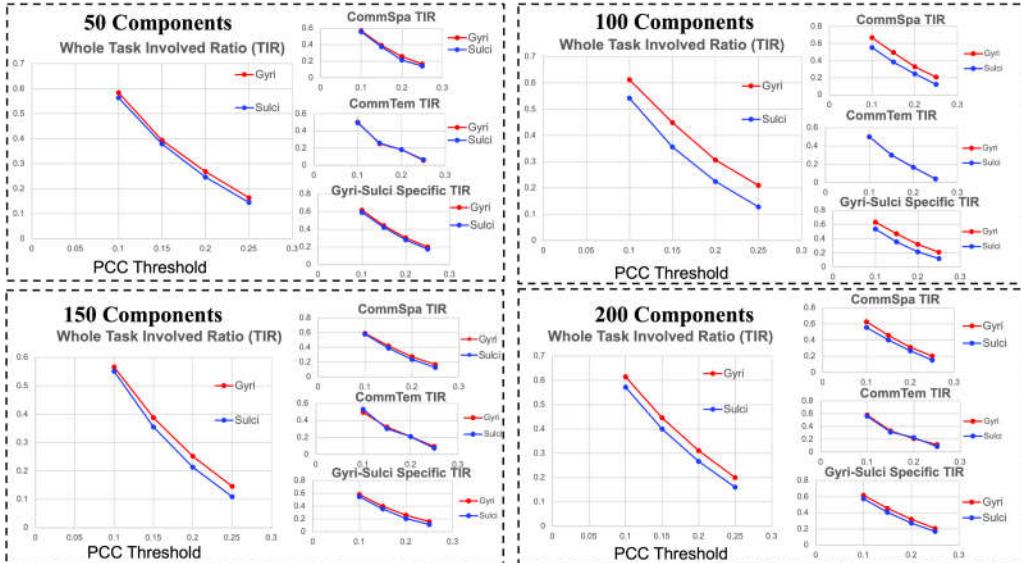


Figure 4: Task Involved Rates. TIR of temporal patterns of gyri and sulci. The temporal patterns are correlated with the task stimulus, and the threshold is set in the range of 0.1, 0.15, 0.2, 0.25. The four plots are the results under different experimental settings of 50, 100, 150, 200 components, where each three subplots are the detailed TIR in different parts.

262 **4.3 Gyri/Sulci/Gyri-Sulci Dominant Network**

263 Besides common FBNs that are derived by enforcing the external constraint, we also achieved a few
 264 FBNs that are categorized as gyri dominant (all the activated voxels belong to gyri), sulci dominant
 265 (all the activated voxels belong to sulci) and gyri-sulci collaborative brain networks (the activated
 266 voxels belong to both gyri and sulci). We display the networks of different categories from randomly
 267 selected 10 subjects in Fig. 5. The results are similar to the common FBNs that gyri dominant FBNs
 268 tend to have large and continuous gyri regions, while sulci dominant ones display scattered and local
 269 distributions. We also analyzed the group-wise ratio between the number of activated brain voxels in
 270 gyri and sulci using different numbers of components in our Twin-Transformer. The results are shown
 271 in Table 2. The gyri ratio is consistently higher than sulci (highlighted in bold). This result indicates
 272 that although there exist sulci dominant BFNs across subjects, the number of activated voxels in gyri
 273 is likely more than that in sulci. In summary, our proposed Twin-Transformer provides a new and
 274 powerful tool to disentangle the different functional roles of gyri and sulci with a new perspective.

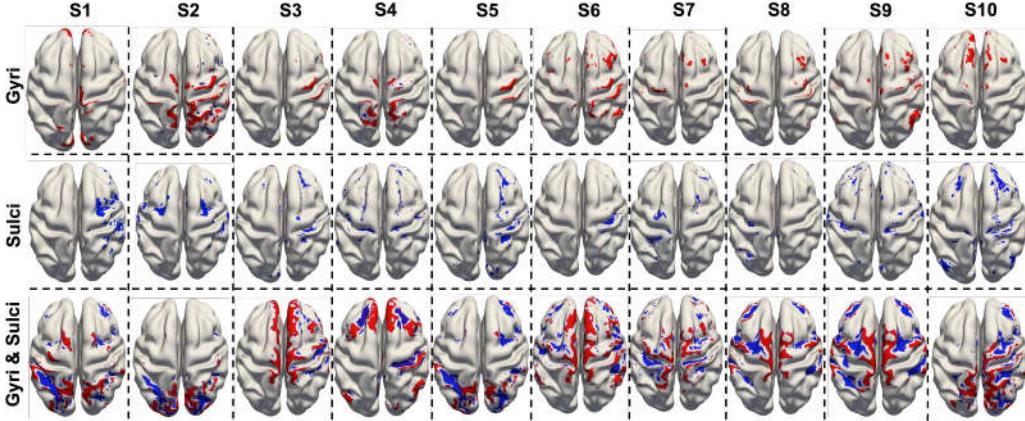


Figure 5: Gyri/Sulci/Gyri-Sulci Dominant Brain Networks. The three rows display of gyri dominant/sulci dominant/gyri-sulci dominant brain networks separately. They are brain functional networks from randomly selected 10 subjects.

275 5 Discussion

276 **Impacts on Brain Science and Artificial Intelligence:** In the brain science field, gyri and sulci
 277 are known to possess different structural, connectional and functional characteristics. However, it
 278 is the first time that our twin-transformer is powerful and accurate enough to differentiate gyri and
 279 sulci into core-periphery networks, which might suggest that the cerebral cortex is segregated into
 280 two fundamentally different functional units of gyri and sulci. This result has profound impacts on
 281 many aspects of basic, cognitive and clinical neuroscience. Core-periphery network phenomena have
 282 been reported in many real-world networked systems such as transportation, social network, financial
 283 networks, and biological neural networks, among others, and our work here revealed and characterized
 284 such core-periphery pattern in a fine-grained manner on cortical gyri and sulci. Given that the graph
 285 structures, e.g., relational graph of CNNs, of artificial neural networks in highly optimized deep
 286 learning models are more similar to those in biological neural networks, it is reasonable to postulate
 287 that the core-periphery network structure discovered in human brains in this work could be potentially
 288 infused into the design of next-generation artificial neural networks in deep learning as a prior
 289 knowledge or meaningful constraint, thus leading to brain-inspired artificial intelligence.

290 **Limitation:** Our work has several potential limitations. a) We simply add the degree from each
 291 activation graph generated by each common spatial component to build the gyri/sulci graph. There
 292 is still some room for building a better gyri/sulci graph. b) We mainly focus on discovering the
 293 relationship between gyri and sulci at this moment, and ignored the intermediate regions on the gyral
 294 wall that is between gyri and sulci. In the near future, we plan to explore the intermediate regions'
 295 roles in the core-periphery brain network system.

296 6 Conclusion

297 In this paper, we proposed a novel data-driven Twin-Transformer framework and applied it to HCP
 298 gray-ordinate tfMRI dataset to characterize the roles of cortical gyri and sulci on the brain functional
 299 networks. With this framework, we can disentangle the spatial and temporal patterns from the brain
 300 signals of gyri and sulci, providing us the possibility to quantitatively analyze the difference between
 301 gyri and sulci. The most important finding in this study is that we identified the core-periphery
 302 relationship between gyri and sulci, as well as the corresponding core-periphery brain networks. Our
 303 results show that core-periphery networks are broadly existing between gyri and sulci across all
 304 subjects. Overall, our proposed Twin-Transformer contributes to a better understanding of the roles
 305 of gyri and sulci in brain architecture, which offers new insight into the design of next-generation
 306 artificial neural networks, brain-inspired AI models, and beyond.

307 **References**

- 308 [1] John P Boyd, William J Fitzgerald, and Robert J Beck. Computing core/periphery structures
309 and permutation tests for social relations data. *Social networks*, 28(2):165–178, 2006.
- 310 [2] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal,
311 Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are
312 few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.
- 313 [3] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and
314 Sergey Zagoruyko. End-to-end object detection with transformers. In *European conference on
315 computer vision*, pages 213–229. Springer, 2020.
- 316 [4] Hanting Chen, Yunhe Wang, Tianyu Guo, Chang Xu, Yiping Deng, Zhenhua Liu, Siwei Ma,
317 Chunjing Xu, Chao Xu, and Wen Gao. Pre-trained image processing transformer. In *Proceedings
318 of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12299–12310,
319 2021.
- 320 [5] Mark Chen, Alec Radford, Rewon Child, Jeffrey Wu, Heewoo Jun, David Luan, and Ilya
321 Sutskever. Generative pretraining from pixels. In *International Conference on Machine
322 Learning*, pages 1691–1703. PMLR, 2020.
- 323 [6] Mihai Cucuringu, Puck Rombach, Sang Hoon Lee, and Mason A Porter. Detection of core-
324 periphery structure in networks using spectral methods and geodesic paths. *European Journal
325 of Applied Mathematics*, 27(6):846–887, 2016.
- 326 [7] Camino De Juan Romero and Victor Borrell. Coevolution of radial glial cells and the cerebral
327 cortex. *Glia*, 63(8):1303–1319, 2015.
- 328 [8] Fan Deng, Xi Jiang, Dajiang Zhu, Tuo Zhang, Kaiming Li, Lei Guo, and Tianming Liu. A
329 functional model of cortical gyri and sulci. *Brain structure and function*, 219(4):1473–1491,
330 2014.
- 331 [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of
332 deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*,
333 2018.
- 334 [10] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai,
335 Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al.
336 An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint
337 arXiv:2010.11929*, 2020.
- 338 [11] Bruce Fischl and Anders M Dale. Measuring the thickness of the human cerebral cortex from
339 magnetic resonance images. *Proceedings of the National Academy of Sciences*, 97(20):11050–
340 11055, 2000.
- 341 [12] Bruce Fischl, Niranjini Rajendran, Evelina Busa, Jean Augustinack, Oliver Hinds, BT Thomas
342 Yeo, Hartmut Mohlberg, Katrin Amunts, and Karl Zilles. Cortical folding patterns and predicting
343 cytoarchitecture. *Cerebral cortex*, 18(8):1973–1980, 2008.
- 344 [13] Caitlyn C Gertz and Arnold R Kriegstein. Neuronal migration dynamics in the developing ferret
345 cortex. *Journal of Neuroscience*, 35(42):14307–14315, 2015.
- 346 [14] Shi Gu, Cedric Huchuan Xia, Rastko Ceric, Tyler M Moore, Ruben C Gur, Raquel E Gur,
347 Theodore D Satterthwaite, and Danielle S Bassett. Unifying the notions of modularity and core-
348 periphery structure in functional brain networks during youth. *Cerebral Cortex*, 30(3):1087–
349 1102, 2020.
- 350 [15] Jeremy Guillon, Mario Chavez, Federico Battiston, Yohan Attal, Valentina La Corte, Michel
351 Thiebaut de Schotten, Bruno Dubois, Denis Schwartz, Olivier Colliot, and Fabrizio de Vico Fal-
352 lani. Disrupted core-periphery structure of multimodal brain networks in alzheimer’s disease.
353 *Network Neuroscience*, 3(2):635–652, 2019.

- 354 [16] Andrew G Haldane and Robert M May. Systemic risk in banking ecosystems. *Nature*,
355 469(7330):351–355, 2011.
- 356 [17] Claus C Hilgetag and Helen Barbas. Developmental mechanics of the primate cerebral cortex.
357 *Anatomy and embryology*, 210(5):411–417, 2005.
- 358 [18] Petter Holme. Core-periphery organization of complex networks. *Physical Review E*,
359 72(4):046111, 2005.
- 360 [19] Christopher J Honey, Jean-Philippe Thivierge, and Olaf Sporns. Can structure predict function
361 in the human brain? *Neuroimage*, 52(3):766–776, 2010.
- 362 [20] Heng Huang, Xintao Hu, Yu Zhao, Milad Makkie, Qinglin Dong, Shijie Zhao, Lei Guo, and
363 Tianming Liu. Modeling task fmri data via deep convolutional autoencoder. *IEEE transactions
364 on medical imaging*, 37(7):1551–1561, 2017.
- 365 [21] Xi Jiang, Tuo Zhang, Shu Zhang, Keith M Kendrick, and Tianming Liu. Fundamental functional
366 differences between gyri and sulci: implications for brain function, cognition, and behavior.
367 *Psychoradiology*, 1(1):23–41, 2021.
- 368 [22] Matthew B Johnson, Peter P Wang, Kutay D Atabay, Elisabeth A Murphy, Ryan N Doan,
369 Jonathan L Hecht, and Christopher A Walsh. Single-cell analysis reveals transcriptional
370 heterogeneity of neural progenitors in human cortex. *Nature neuroscience*, 18(5):637–646,
371 2015.
- 372 [23] Gang Li, Tianming Liu, Dong Ni, Weili Lin, John H Gilmore, and Dinggang Shen. Spatiotem-
373 poral patterns of cortical fiber density in developing infants, and their relationship with cortical
374 thickness. *Human brain mapping*, 36(12):5183–5195, 2015.
- 375 [24] Qing Li, Xia Wu, and Tianming Liu. Differentiable neural architecture search for optimal
376 spatial/temporal brain function network decomposition. *Medical Image Analysis*, 69:101974,
377 2021.
- 378 [25] Huan Liu, Shu Zhang, Xi Jiang, Tuo Zhang, Heng Huang, Fangfei Ge, Lin Zhao, Xiao Li, Xintao
379 Hu, Junwei Han, et al. The cerebral cortex is bisectinally segregated into two fundamentally
380 different functional units of gyri and sulci. *Cerebral Cortex*, 29(10):4238–4252, 2019.
- 381 [26] Huan Liu, Shu Zhang, Xi Jiang, Tuo Zhang, Heng Huang, Fangfei Ge, Lin Zhao, Xiao Li, Xintao
382 Hu, Junwei Han, et al. The cerebral cortex is bisectinally segregated into two fundamentally
383 different functional units of gyri and sulci. *Cerebral Cortex*, 29(10):4238–4252, 2019.
- 384 [27] Huan Liu, Shu Zhang, Xi Jiang, Tuo Zhang, Heng Huang, Fangfei Ge, Lin Zhao, Xiao Li, Xintao
385 Hu, Junwei Han, et al. The cerebral cortex is bisectinally segregated into two fundamentally
386 different functional units of gyri and sulci. *Cerebral Cortex*, 29(10):4238–4252, 2019.
- 387 [28] Huan Liu, Shu Zhang, Xi Jiang, Tuo Zhang, Heng Huang, Fangfei Ge, Lin Zhao, Xiao Li, Xintao
388 Hu, Junwei Han, et al. The cerebral cortex is bisectinally segregated into two fundamentally
389 different functional units of gyri and sulci. *Cerebral Cortex*, 29(10):4238–4252, 2019.
- 390 [29] Jinglei Lv, Xi Jiang, Xiang Li, Daqiang Zhu, Hanbo Chen, Tuo Zhang, Shu Zhang, Xintao
391 Hu, Junwei Han, Heng Huang, et al. Sparse representation of whole-brain fmri signals for
392 identification of functional networks. *Medical image analysis*, 20(1):112–134, 2015.
- 393 [30] Jinglei Lv, Xi Jiang, Xiang Li, Daqiang Zhu, Shu Zhang, Shijie Zhao, Hanbo Chen, Tuo Zhang,
394 Xintao Hu, Junwei Han, et al. Holistic atlases of functional networks and interactions reveal
395 reciprocal organizational architecture of cortical function. *IEEE Transactions on Biomedical
396 Engineering*, 62(4):1120–1131, 2014.
- 397 [31] Jingxin Nie, Lei Guo, Kaiming Li, Yonghua Wang, Guojun Chen, Longchuan Li, Hanbo Chen,
398 Fan Deng, Xi Jiang, Tuo Zhang, et al. Axonal fiber terminations concentrate on gyri. *Cerebral
399 cortex*, 22(12):2831–2839, 2012.

- 400 [32] Jonas Richiardi, Andre Altmann, Anna-Clare Milazzo, Catie Chang, M Mallar Chakravarty,
 401 Tobias Banaschewski, Gareth J Barker, Arun LW Bokde, Uli Bromberg, Christian Büchel,
 402 et al. Correlated gene expression supports synchronous activity in brain networks. *Science*,
 403 348(6240):1241–1244, 2015.
- 404 [33] M Puck Rombach, Mason A Porter, James H Fowler, and Peter J Mucha. Core-periphery
 405 structure in networks. *SIAM Journal on Applied mathematics*, 74(1):167–190, 2014.
- 406 [34] Camille Roth, Soong Moon Kang, Michael Batty, and Marc Barthelemy. A long-time limit for
 407 world subway networks. *Journal of The Royal Society Interface*, 9(75):2540–2550, 2012.
- 408 [35] Vanessa Troiani, Marisa A Patti, and Kayleigh Adamson. The use of the orbitofrontal h-sulcus
 409 as a reference frame for value signals. *European Journal of Neuroscience*, 51(9):1928–1943,
 410 2020.
- 411 [36] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez,
 412 Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information
 413 processing systems*, 30, 2017.
- 414 [37] Shimin Yang, Zhongbo Zhao, Han Cui, Tuo Zhang, Lin Zhao, Zhibin He, Huan Liu, Lei Guo,
 415 Tianming Liu, Benjamin Becker, et al. Temporal variability of cortical gyral-sulcal resting state
 416 functional activity correlates with fluid intelligence. *Frontiers in neural circuits*, page 36, 2019.
- 417 [38] Sixiao Zheng, Jiachen Lu, Hengshuang Zhao, Xiatian Zhu, Zekun Luo, Yabiao Wang, Yanwei
 418 Fu, Jianfeng Feng, Tao Xiang, Philip HS Torr, et al. Rethinking semantic segmentation
 419 from a sequence-to-sequence perspective with transformers. In *Proceedings of the IEEE/CVF
 420 conference on computer vision and pattern recognition*, pages 6881–6890, 2021.
- 421 [39] Luowei Zhou, Yingbo Zhou, Jason J Corso, Richard Socher, and Caiming Xiong. End-to-end
 422 dense video captioning with masked transformer. In *Proceedings of the IEEE Conference on
 423 Computer Vision and Pattern Recognition*, pages 8739–8748, 2018.

424 Checklist

425 The checklist follows the references. Please read the checklist guidelines carefully for information on
 426 how to answer these questions. For each question, change the default [TODO] to [Yes] , [No] , or
 427 [N/A] . You are strongly encouraged to include a **justification to your answer**, either by referencing
 428 the appropriate section of your paper or providing a brief inline description. For example:

- 429 • Did you include the license to the code and datasets? [Yes] See Section ?? . Yes
 430 • Did you include the license to the code and datasets? [No] The code and the data are
 431 proprietary.
 432 • Did you include the license to the code and datasets? [N/A]

433 Please do not modify the questions and only use the provided macros for your answers. Note that the
 434 Checklist section does not count towards the page limit. In your paper, please delete this instructions
 435 block and only keep the Checklist section heading above along with the questions/answers below.

- 436 1. For all authors...
 - 437 (a) Do the main claims made in the abstract and introduction accurately reflect the paper’s
 438 contributions and scope? [TODO] Yes
 - 439 (b) Did you describe the limitations of your work? [TODO] Yes
 - 440 (c) Did you discuss any potential negative societal impacts of your work? [TODO] No
 - 441 (d) Have you read the ethics review guidelines and ensured that your paper conforms to
 442 them? [TODO] Yes
- 443 2. If you are including theoretical results...
 - 444 (a) Did you state the full set of assumptions of all theoretical results? [TODO] Yes
 - 445 (b) Did you include complete proofs of all theoretical results? [TODO] Yes