模拟视皮层功能柱的动作识别神经网络

朱后颖 陈新欣 方璐璐 严睿恒 刘海华[通信作者]

中南民族大学 生物医学工程学院，武汉，430074；

摘要：大脑视皮层存在大量具有相似功能神经元组成的功能柱，其作为处理视觉信息的基本单元，在完成各种视觉任务中发挥重要作用.为此，提出了模拟视皮层方位功能柱结构的卷积神经网络模型（FCNet），并将其应用于人体动作识别任务.该网络模型模拟视皮层功能柱生物学特性，利用三维时空 Gabor 滤波器构建了计算功能柱；以 CNN 网络为骨架，计算功能柱为卷积核组，构建了动作识别神经网络；该网络采用直接分裂式前馈连接方式，提取视频中的时空特征，完成动作识别任务.通过在 KTH 和 UCF101 等公共动作识别数据集上进行实验，其结果表明，FCNet 在动作识别准确率和效率方面显著优于其他卷积神经网络模型.其中，在 KTH 数据集上的分类准确率高达 92.93%，在 UCF101 数据集上分类准确率达到 90.04%，而与其他模型相比减少大量的参数量和计算代价.

关键词：3D Gabor 滤波器；方向选择性；视觉功能柱；动作识别

An action recognition neural network that simulates the function column structure of the visual cortex

ZHU Houying, CHEN Xinxin, FANG Lulu, YAN Ruiheng, LIU HaihuaCorresponding author

College of Biomedical Engineering, South-Central Minzu University, Wuhan China，430074；

Abstract：The cerebral visual cortex contains a large number of function columns composed of neurons with similar functions. These columns serve as fundamental units for visual information processing and play crucial roles in performing various visual tasks. This paper proposes a convolutional neural network model (FCNet) that mimics the orientation function column structure of the visual cortex is proposed and applied to the action recognition task. This network model emulates the biological characteristics of the functional columns in the visual cortex. It constructs computational function columns using three-dimensional spatio-temporal Gabor filters. With the CNN network as the framework and the computational functional columns as the convolutional kernel groups, an action recognition neural network is built. This network employs a direct split-type feed-forward connection method to extract spatio-temporal features from videos and complete the action recognition task. Evaluated on public action recognition datasets KTH and UCF101, experimental results demonstrate that FCNet significantly outperforms other CNN models in both recognition accuracy and computational efficiency. Specifically, it achieves classification accuracies of 92.93% on KTH and 90.04% on UCF101, while substantially reducing the number of parameters and computational costs compared to other models.

Keywords：3D Gabor filter；orientation selectivity；visual function column；action recognition

DOI:10.69979/3029-2808.25.06.055

人体动作识别是视频分析的主要内容，在人机交互、智能监测、行为分析、医学康复等方面有着广阔的应用前景.因此，视频人体动作识别也是计算机视觉领域研究的热点之一，吸引大量学者参与人体动作的研究，并取得了一些研究成果. 然而，由于人体动作的复杂性、动作所在环境的多样性，使得人体动作识别仍存在许多有待解决的问题.

针对人体动作识别中所存在的挑战问题，学者们利用深度学习开展人体动作识别研究，提出了大量脑启发式动作识别方法[1-4].这些方法主要分为两类:1）模拟视皮层经典感受野的 3D 卷积神经网络（3D-CNN）方法[5-7]；2）模拟视皮层信息加工双通道理论的双流神经网络方法[8-13].双流方法利用了视皮层腹侧通路和背侧通路分别加工空间和运动信息的特性，分别处理视频图像（空间）和光流（运动）信息，从而实现人体动作识别.3D-CNN 的方法不仅利用了视觉神经元的经典感受野属性，而且利用了其时空相关性提取特征，有利于动作识别.然而，由于 3D-CNN 计算复杂度较高，因此大量学者开展了减少计算量、提升识别性能的扩展性研究[14-20].然而，这些扩展性的方法使 3D-CNN 的结构越来越复杂.

视觉神经生理学研究中，Hubel 和 Wiesel 发现神经细胞对不同类型的光刺激其反应是不同的[21]，即神经元

对一定朝向（或方位）的光栅具有强烈反应，而对偏离该朝向的光栅反应较弱，且随着偏离度越大，响应衰减越强. 且具有此类特性的神经元在视皮层形成了生物学上的功能柱[22]. 为此，本文利用视皮层神经元的这种方向选择性，构建以时空相关 Gabor 滤波器为基础的方位功能柱，建立基于计算功能柱的卷积神经网络模型，实现对人体动作的识别，推进类脑计算的研究.

1 三维功能柱卷积神经网络

1.1 功能柱计算模型

Hubel 和 Wiesel 的研究发现，视皮层神经元对不同朝向（或方位）的刺激具有一定的“偏好”. 而 Daugman 根据神经元的这种生物学属性，利用 2D-Gabor 函数对其响应进行了描述[23]，反映了其对刺激的频率、方向选择性. 在此，为了处理视频中的时空信息，将 2D-Gabor 滤波器扩展为 3D-Gabor 滤波器，其数学表达为:

$$G_{\sigma,v,\theta,\phi}(x,y,t)=\frac{\gamma}{2\pi\sigma^2}\exp\left[\frac{-\left((\bar{x}-v_c t)^2+\gamma^2\bar{y}^2\right)}{2\sigma^2}\right]\times \cos\left(\frac{2\pi}{\lambda}(\bar{x}-vt)+\phi\right)\times\frac{1}{\sqrt{2\pi}\tau}\exp\left[-\frac{(t-\mu_t)^2}{2\tau^2}\right]U(t), \quad (1)$$

$$\bar{x}=x\cos(\theta)+y\sin(\theta),\bar{y}=-x\sin(\theta)+y\cos(\theta), \quad (2)$$

其中 $U(t)$为阶跃函数，v 为运动速度，v_c为空间高斯包的水平轴移动速度. 当$v_c=0$ 时，空间高斯包的中心是静止的. 运动速度 v 和波长 λ 遵循以下约束：

$$\lambda=\lambda_0\sqrt{1+v^2}, \quad (3)$$

其中参数 λ_0为常数. 由此构建了运动速度与空间频率间的关系. 当 $\lambda_0=2$，$\sigma=1.12\sqrt{1+v^2}$. 时间高斯均值 μ_t和标准差 τ 根据生理学实验数据设置为：$\mu_t=1.75$，$\tau=2.75$.

由式(1)可以看出，3D-Gabor 函数很好地模拟了视皮层神经元的频率选择性和方向选择性. 为此，根据视皮层神经元的方向选择性，建立具有方向选择性的计算功能柱$FG_\sigma(x,y,t)$，其数学表达为：

$$FG_\sigma(x,y,t)=\begin{bmatrix}G_{\sigma,\theta_1}(x,y,t)\\G_{\sigma,\theta_2}(x,y,t)\\\vdots\\G_{\sigma,\theta_n}(x,y,t)\end{bmatrix}, \quad (4)$$

其中 σ 为该功能柱的空间频率，θ_i为第 i 个 Gabor 核的“偏好”方向，n 为功能柱的密度，即功能柱所包含的 Gabor 核数量.

1.2 计算功能柱的设计

根据式(4)的功能柱计算模型，每个方位功能柱中包含了 n 个 Gabor 函数，所有 Gabor 核具有相同的空间频率，而每个 Gabor 核具有不同的方向. 根据该功能柱计算模型，计算功能柱的设计是以该模型为基础的卷积操作，其数学表达为：

$$\text{Output}=\text{ConvFG}_\sigma(\text{Input}), \quad (5)$$

其中 Input 和 Output 为计算功能柱的输入和输出. ConvFGσ 是以功能柱计算模型为核的卷积操作，从而构建了基本的计算功能柱.

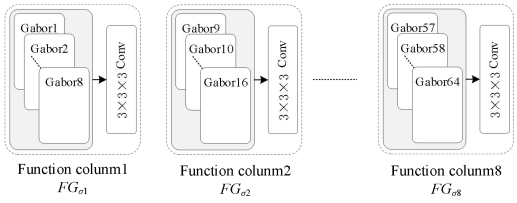

图 1 三维方位功能柱

Fig.1 Three dimensional orientation function columns

为了进一步提高功能柱提取时空特征的能力，在基本计算功能柱的基础上，级联一个通用的卷积操作，从而构建了扩展的计算功能柱，其数学表达为：

$$\text{Output}=\text{Conv}\left(\text{ConvFG}_\sigma(\text{Input})\right), \quad (6)$$

其中 Conv 是核大小为 3×3×3 的 3D 卷积。该计算功能柱仍用 ConvFGσ 表述。

此外，计算功能柱的设计还依赖于神经网络结构，即在每个网络层中，计算功能柱的数量由网络层的通道数和功能柱的密度决定。假设网络层的通道为 C，功能柱的密度为 n，则该网络层有 C/n 个功能柱，将该网络层称为功能柱网络层. 例如，64 个通道的功能柱网络层设计如图 1 所示。其中功能柱的密度为 8，该网络层由 8 个网络功能柱组成，每个功能柱具有不同的空间频率。基于该设计，该网络层需要学习 8 个空间频率参数 σ_j，64 个方位参数 θ_i. 为了提高每个功能柱对局部时空特征的捕捉能力，将式(1)定义的 3D-Gabor 核中参数 ϕ、γ 释放为可学习参数.

1.3 基于计算功能柱的网络模型

基于计算功能柱的神经网络模型(FCNet)的设计根据经典 3D-CNN 网络结构（C3D）为基础的. 基本的设计思想是，用计算功能柱网络层替换原 C3D 网络中的卷积层，如图 2 所示。图 2 分别给出了针对低空间分辨率和高空间分辨率视频的网络结构。在本文中，图 2(a)和(b)

的结构分别应用于 KTH 和 UCF101 人体动作识别数据集. 但无论是哪种结构，建立的 FCNet 网络结构只用计算功能柱网络层替代原网络结构中的前 3 个卷积层，而保留原网络后续结构. 根据生物学的研究成果，将深网络层中功能柱密度设置为浅网络层中功能柱密度的一半。假设第 i 层中功能柱的密度为 ni，则第 i+1 层中功能柱的密度为 ni/2. 据此，FCNet 中的第一到第三网络层中功能柱密度分别为 8、4、2，其相应的功能柱数量为 8、32、128. 同时，在每个计算功能柱之后，增加 BN 层，以此加强信息处理.

由于计算功能柱将网络层分为多个组，这就导致网络层之间功能柱如何连接的问题。如果采用通用的卷积层连接方式，既不能充分发挥功能柱的作用，也会增加计算量。为此，FCNet 根据分组卷积的思想处理该问题. 分组卷积是深度学习的一种技术[24]，旨在通过将通道预定地分成多个组别，每组间独立地进行卷积操作，从而构建稀疏网络，提高模型的计算效率和特征提取能力. 然而，由于 FCNet 中网络层之间功能柱的密度和功能柱的数量是不同的，因此不能直接采用通用的分组卷积技术。为此，FCNet 根据网络层中功能柱的密度和功能柱的数量的不同，采用直接分裂的方式构建功能柱网络层之间的连接，如图 3 所示.

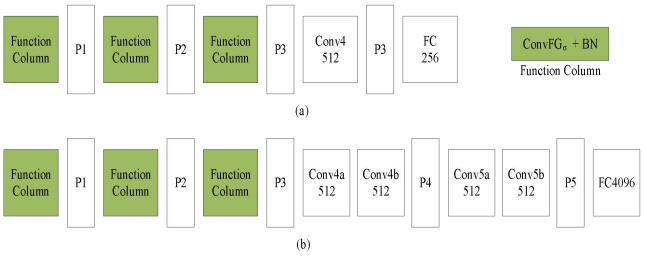

图 2 FCNet 网络架构

Fig.2 FCNet network architecture

直接分裂的方式是将前一层中的 1 个功能柱与下一层中的多若干个功能柱建立前馈连接。假设前一网络层中功能柱（Previous layer column）的密度为 4，则将该功能柱在下一层中分裂为 2 个功能柱（Next layer column），其密度为 2，分裂后的功能柱只与该共嗯组建立连接，而不与其它功能柱建立连接，从而形成了稀疏网络连接，如图 3 所示. 根据如图 2 所示的 FCNet 的结构，前一网络层通道数是下一层网络层通道数的 1/2 倍，而功能柱的密度与下一层功能柱大小一样，或者是其 2 倍，则该网络层中的 1 个功能柱将在下一层分裂为 2 个或 4 个功能柱。由此可见，根据功能柱密度的不同，FCNet 网络结构的配置有所不同。

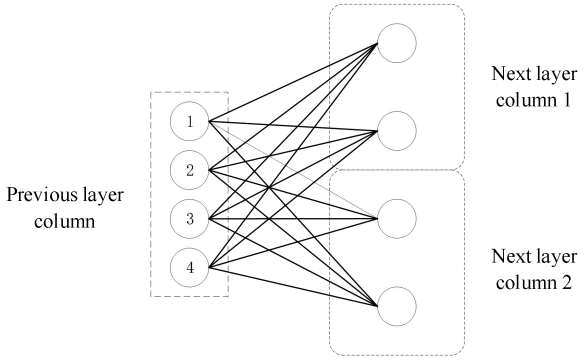

图 3 功能柱直接分裂式前馈连接

Fig.3 Direct split-type feed-forward connection of function columns

2 实验结果与分析

为了评估建议的 FCNet 的性能，将其在公开的 KTH 和 UCF101 经典动作识别数据集进行实验. KTH 数据集由步行、慢跑、跑步、拳击、挥手和拍手等 6 种类型的动作组成. 这些视频由 25 个受试者，使用静态摄像机录制. 该数据库不仅包含大尺度变化的视频样本，还包含衣服颜色样式变化和背景光照变化较大的视频样本.

UCF101 数据源自 YouTube 视频网站，其中包括了 101 种独特的运动类型，例如潜泳、滑雪、骑马等. 整个 UCF101 数据包共有 13320 个视频样本，其样本的长短与清晰程度各异，且在不同光线条件下以不同的视角录制而成。所有视频样本都被标注了相应的动作标签.

此外，所有实验均在 Centos7 系统上完成. 该系统配备了 Intel(R) Xeon(R) Gold 5118 CPU @2.3.GHz 和 NVIDIA GeForce-3090-24GB 的 GPU；FCNet 在该平台上由 Anaconda4.5.4，Python3.8.9，Pytorch1.2.0 构建.

首先，在 KTH 数据集上对 FCNet 网络模型进行实验，以评估其性能. 实验中从数据集中随机抽取 16 名动作执行者视频作为训练集，而剩余 9 名动作执行者视频为测试集. 对网络训练时，采用 Adam 优化器，初始学习率为 0.001，经过 10 个 epoch 学习率降低 10 倍. 表 1 给出了不同配置 FCNet 的性能，其中 FCNet-1、FCNet-2、FCNet-3 为 C3D 的前 1、2、3 卷积层分别被功能柱网络层替代的网络模型，括号中的数字表示相应网络层中功能柱的密度. 从表 1 中的结果可以看出，无论是哪种网络结构，FCNet 的性能都高于 C3D 的性能；第一到第三网络层中功能柱大小不变的 FCNet 网络性能低于功能柱密度

逐渐减少的网络性能. 如 FCNet-2 (8-4)的性能(92.54%)比 FCNet-2(8-8)的性能(91.39%)高，FCNet-3(8-4-2)的性能(91.11%)比 FCNet-3(8-8-8)的性能(92.93%)高. 该结果与神经生物学的研究成果是相一致. 也就是说，FCNet 网络层中的功能柱密度逐渐减少，能够更好地处理空信息. 此外，从 FCNet-1 到 FCNet-3，其分类准确率不仅逐步提升，而且计算量逐步降低，其中 FCNet-3 的计算量(1.28G)远低于 C3D(5.96G)计算量.

表 1 直接分组式层连接功能柱在 KTH 数据集上分类准确率

Tab.1 Classification accuracy of directly grouped layer-connected function columns on the KTH dataset

方法	分类准确率/%	参数量/M	计算量/GFLOPs
C3D	88.23	21.42	5.96
FCNet-1(8)	91.99	20.56	5.64
FCNet-2(8-8)	91.39	23.68	2.19
FCNet-2(8-4)	92.54	23.89	2.34
FCNet-3(8-8-8)	91.11	18.39	1.00
FCNet-3(8-4-2)	92.93	21.14	1.28

其次，在更复杂的 UCF101 数据集上对 FCNet 网络模型进行实验，进一步评估其性能. 表 2 列出了在最优配置条件下的实验结果，即使用功能柱密度逐渐减少的 FCNet 的实验结果. 从实验结果可以发现，FCNet-1 到 FCNet-3 的分类准确率虽然若有降低，但都高于 C3D 的分类准确率，且下降幅度非常有限；更重要的是 FCNet-1 到 FCNet-3 的计算量逐渐减少，其中 FCNet-3 的计算量(27.17G)仅为 FCNet-1 计算量(37.51G)的 72%，是 C3D 计算量(38.83G)的 70%. 这表明建议的 FCNet 在复杂数据集上既有有较好的分类性能，也有较高的计算效率.

表 2 直接分组式层连接功能柱在 UCF101 数据集上分类准确率

Tab.2 Classification accuracy of directly grouped layer-connected function columns on the UCF101 dataset

方法	分类准确率（%）	参数量（M）	计算量（GFLOPs）
C3D	85.20	79.42	38.83
FCNet-1(8)	90.56	79.40	37.51
FCNet-2(8-4)	90.37	79.23	32.34
FCNet-3(8-4-2)	90.04	78.53	27.17

为了验证建议的 FCNet 网络模型的有效性，表 3 给出了其与当前最先进的网络模型的性能比较. 从计算效率来看，建议的 FCNet-3 网络模型的计算代价是最少的，仅为改进的 C3D(T-C3D)[19]计算代价的一半；从分类准确率来看，FCNet-3 的准确率仅次于 TSN[13]和 VIMPAC[25]，但这两个网络模型是采用预训练或编码-解码后所获得的结果，而 FCNet 是直接训练后所取得的结果；从参数量来看，FCNet-3 的参数量似乎是比较高的，与原始的 C3D 相当，但该参数量为推理阶段的计算参数量，而不是训练参数量. FCNet-3 的训练参数量与 I3D 的参数量相当，这表明了 FCNet 模型有较高的训练效率。

表 3 不同研究方法在 UCF101 数据集上的分类准确率

Tab.3 Classification accuracy of different research methods on the UCF101 dataset

方法	分类准确率（%）	参数量（M）	计算量（GFLOPs）
C3D	85.2	79.42	38.8
Two-stream	88.0	50	80
I3D	84.5	12.3	108
TSN	94.2	50	80
R3D-18	86.2	33	56
T-C3D	89.4	85	50
R(2+1)D	78.7	-	-
VIMPAC	92.7	-	-
FCNet-3	90.04	78.53	27.17

具体而言，双流网络或基于双流网络架构改进的 I3D、R3D-18 网络等模型[9-10]，能结合光流提取联合时空特征，对短时动作有较好的分类效果. 然而，由于光流的预计算等因素影响，导致 TSN 依赖大规模的视频数据集预训练才能发挥其性能优势；I3D 网络在空间流中进行了时空联合建模，导致较高的计算量. 基于 C3D 网络模型改进的 T-C3D、R(2+1)D 等网络模型[20]，尽管增加了多尺度变化，加强长时序感知，相比于 C3D 的分类准确率有所提升，但参数效率和计算效率有所降低. 综合而言，本文模拟视皮层功能柱所提出的 FCNet-3 网络模型，能更好地处理视频中的时空信息，不仅分类性能较好，而且计算量效率最高.

3 结论

本文利用 3D-Gabor 滤波器模拟视皮层简单细胞的方向选择性和特定的空间频率的敏感性，提出了功能柱计算模型，并利用该计算模型建立了计算功能柱卷积层，以此替换三维卷积神经网络中的传统卷积层，构建了基于计算功能柱的网络模型(FCNet). 在该网络模型中，功能柱的密度随着网络深度加深而减少，网络层之间的功能柱采用直接分裂的方式相连，从而提高了对时空信息处理的能力. 建议的 FCNet 在 KTH 和 UCF101 数据集上的实验结果表明，其在动作识别任务上的分类准确率优于

基于三维卷积的神经网络模型；基于 Gabor 滤波器的功能柱计算模型以及层级连接方式，使该网络与其它网络模型相比具有更少的计算量和更少的训练参数量. 由于 FCNet 采用直接分裂的方式构建网络层间的连接，在一定程度上限制的网络模型的优势，在此后的研究中，可探讨其它连接方式，以提升 FCNet 在动作识别中的能力，检验方法的鲁棒性.

参考文献

[1] BARSHOOI A H, AMIRKHANI A. A novel data augmentation based on Gabor filter and convolutional deep learning for improving the classification of COVID-19 chest X-Ray images[J]. Biomedical Signal Processing and Control, 2022, 72: 103326.

[2] ALEKSEEV A, BOBE A. GaborNet: Gabor filters with learnable parameters in deep convolutional neural network[C]//IEEE. 2019 International Conference on Engineering and Telecommunication. Dolgoprudny: IEEE, 2019: 1-4.

[3] RAJPUT S S, CHOI Y. Handwritten digit recognition using Convolution Neural Networks[C]// IEEE. 2022 IEEE 12th Annual Computing and Communication Workshop and Conference. Las Vegas: IEEE, 2022: 0163-0168.

[4] LUAN S, CHEN C, ZHANG B, et al. Gabor convolutional networks[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4357-4366.

[5] YUAN Y, WANG L N, ZHONG G, et al. Adaptive Gabor convolutional networks[J]. Pattern Recognition, 2022, 124: 108495.

[6] FEICHTENHOFER C, FAN H, MALIK J, et al. SlowFast networks for video recognition[C]//IEEE. 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 6201-6211.

[7] TRAN D, BOURDEV L, FERGUS R, et al. Learning spatiotemporal features with 3d convolutional networks[C]//IEEE. 2015 IEEE International Conference on Computer Vision. Santiago:IEEE, 2015: 4489-4497.

[8] SIMONYAN K, ZISSERMAN A. Two-stream convolutional networks for action recognition in videos[C]//NIPS. 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2014: 568-576.

[9] CARREIRA J, ZISSERMAN A. Quo vadis, action recognition? A new model and the kinetics dataset[C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu:IEEE, 2017: 6299-6308.

[10] REHMAN Y A U, GAO Y, SHEN J, et al. Federated self-supervised learning for video understanding[C]// Springer. European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 506-522.

[11] LIN J, GAN C, HAN S. TSM: Temporal shift module for efficient video understanding[C]//IEEE. IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 7083-7093.

[12] QIU Z, YAO T, NGO C W, et al. Learning spatio-temporal representation with local and global diffusion[C]//IEEE. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 12056-12065.

[13] WANG L, XIONG Y, WANG Z, et al. Temporal segment networks: Towards good practices for deep action recognition[C]//Springer. European conference on computer vision. Cham: Springer International Publishing, 2016: 20-36.

[14] FEICHTENHOFER C. X3D: Expanding architectures for efficient video recognition[C]//IEEE. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 203-213.

[15] TAYLOR G W, FERGUS R, LECUN Y, et al. Convolutional learning of spatio-temporal features[C]// Springer. Computer Vision-ECCV 2010: 11th European Conference on Computer Vision. Berlin Heidelberg: Springer, 2010: 140-153.

[16] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//IEEE. IEEE Conference on Computer Vision and Pa

ttern Recognition. Honolulu: IEEE, 2017: 4700-4708.
[17] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv:2017, 1704.04861.
[18] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 6848-6856.
[19] LIU K, LIU W, GAN C, et al. T-C3D: Temporal convolutional 3D network for real-time action recognition[C]//AAAI. The Thirty-Second AAAI Conference on Artificial Intelligence. New Orleans: AAAI, 2018, 32(1).
[20] PAN T, SONG Y, YANG T, et al. Videomoco: Contrastive video representation learning with temporally adversarial examples[C]//IEEE. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 11205-11214.
[21] HUBEL D H, WIESEL T N. Sequence regularity and geometry of orientation columns in the monkey striate cortex[J]. Journal of Comparative Neurology, 1974, 158(3): 267-293.
[22] PETKOV N, SUBRAMANIAN E. Motion detection, noise reduction, texture suppression, and contour enhancement by spatiotemporal Gabor filters with surround inhibition[J]. Biological Cybernetics, 2007, 97: 423-439.
[23] DAUGMAN J G. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 2002, 36(7): 1169-1179.
[24] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
[25] TAN H, LEI J, WOLF T, et al. Vimpac: Video pre-training via masked token prediction and contrastive learning[J].arXiv:2021, 2106.11250.

作者简介:朱后颖(2000-),男,硕士,研究方向:视觉认知计算。
通信作者:刘海华(1967-),男,教授,博士,研究方向:视觉认知计算。
基金项目:湖北省自然科学基金资助项目(61773409)