前额叶皮层启发的 Transformer 模型应用及其进展

潘雨辰 [1, 2] 贾克斌 [2] 张铁林 [1, 3, 4]

摘 要 聚焦于生物结构与类脑智能的交叉研究方向, 探讨前额叶皮层的结构及其认知功能对人工智能领域 Transformer 模型的启发. 前额叶皮层在认知控制和决策制定中扮演着关键角色. 首先介绍前额叶皮层的注意力机制、生物编码、多感觉融合等相关生物研究进展, 然后探讨这些生物机制如何启发新型的类脑 Transformer 架构, 重点提升其在自注意力、位置编码、多模态整合等方面的生物合理性与计算高效性. 最后, 总结前额叶皮层启发的类脑新模型, 在支持多类型神经网络组合、多领域应用、世界模型构建等方面的发展与潜力, 为生物和人工智能两大领域之间交叉融合构建桥梁.

关键词 生物结构, 类脑智能, 前额叶皮层, Transformer, 世界模型

引用格式 潘雨辰, 贾克斌, 张铁林. 前额叶皮层启发的 Transformer 模型应用及其进展. 自动化学报, 2025, **51**(7): 1403−1422

DOI 10.16383/j.aas.c240538 **CSTR** 32138.14.j.aas.c240538

The Application and Progress of Prefrontal Cortex-inspired Transformer Model

PAN Yu-Chen[1, 2] JIA Ke-Bin[2] ZHANG Tie-Lin[1, 3, 4]

Abstract This article focuses on the integration of biological network connectomes and brain-inspired intelligence, exploring how the structure and cognitive functions of the prefrontal cortex can inspire transformer models in the field of artificial intelligence. The prefrontal cortex plays a critical role in cognitive control and decision-making. Firstly, the article introduces some advancements of biological research related to the prefrontal cortex′s attention mechanisms, biological encoding, and multisensory integration. Then, it discusses how these biological mechanisms can inspire novel brain-inspired transformer architectures, with a focus on enhancing their biological plausibility and computational efficiency in self-attention, positional encoding, and multimodal integration. Finally, it summarizes the development and potential of new brain-inspired models influenced by the prefrontal cortex, highlighting their support for the integration of various neural network types, multi-domain applications, and the construction of world models in reinforcement learning, thereby building a bridge between the fields of biology and artificial intelligence.

Key words Biological structures, brain-inspired intelligence, prefrontal cortex, transformer, world model

Citation Pan Yu-Chen, Jia Ke-Bin, Zhang Tie-Lin. The application and progress of prefrontal cortex-inspired transformer model. *Acta Automatica Sinica*, 2025, **51**(7): 1403−1422

近年来, 随着生物脑图谱检测技术的不断发展, 前额叶皮层 (Prefrontal cortex, PFC) 在生物神经科学领域的研究取得显著进展, 人们对 PFC 的结构与功能有着更深刻的认识. PFC 是大脑中负责高层次认知功能的核心区域, 具有决策制定、工作记忆和认知控制等高级认知功能. 通过对 PFC 的深入研究, 科学家们能够更全面地理解人类行为背后的神经基础, 并以此探索出相关的神经疾病治疗方法. 与此同步, 人工智能 (Artificial intelligence, AI) 领域也在蓬勃发展, 特别是 2017 年 Transformer 模型的提出为 AI 领域带来深刻的变革. Transformer 模型作为一种基于自注意力机制的深度学习架构, 已经在自然语言处理、计算机视觉等多个应用领域中展现出卓越的效果. 有意思的是, 人工智能领域的大模型架构是以 Transformer 为基本单元堆叠而成, 与之相似, 生物智能领域的 PFC 架构也是以类似的皮质柱结构为基本单元复用实现, 两者的殊途同归给研究者带来更大的信心和兴

收稿日期 2024-07-30 录用日期 2024-12-13
Manuscript received July 30, 2024; accepted December 13, 2024
北京市科技新星 (20230484369), 上海市市级科技重大专项 (2021SHZDZX), 中国科学院青促会基金, 多模态人工智能系统全国重点实验室开放课题基金资助
Supported by Beijing Nova Program (20230484369), Shanghai Municipal Science and Technology Major Project (2021SHZDZX), Youth Innovation Promotion Association of Chinese Academy of Sciences, and Open Projects Program of State Key Laboratory of Multimodal Artificial Intelligence Systems
本文责任编委 雷柏英
Recommended by Associate Editor LEI Bai-Ying
1. 中国科学院脑科学与智能技术卓越创新中心 上海 200031 2. 北京工业大学信息科学技术学院 北京 100124 3. 中国科学院大学 北京 100049 4. 中国科学院自动化研究所 北京 100190
1. Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031 2. School of Information Science and Technology, Beijing University of Technology, Beijing 100124 3. University of Chinese Academy of Sciences, Beijing 100049 4. Institute of Automation, Chinese Academy of Sciences, Beijing 100190

趣, 探讨更多的可能路径去融合、启发、促进两者的向前发展.

随着对 PFC 结构与功能研究的逐步深入, 以及 Transformer 模型在 AI 领域的广泛成功应用, 把生物理论与人工智能相结合的跨学科研究必将是进一步促进新一代人工智能模型发展的关键动力. 通过将 PFC 的生物机制与 Transformer 模型架构相结合, 可以实现更符合"人类认知特征"的 AI 模型, 这种结合不仅有助于提升 AI 模型在处理复杂任务时的性能, 而且还可以搭建起生物脑智能和人工智能之间的桥梁, 为 AI for Neuroscience 提供新视角.

如图 1 所示, 本文首先在第 1 节中对 PFC 进行生物层面的分析, 对它所具有的功能以及目前在类脑智能领域的应用进行介绍. 在第 2 节中通过三个部分进一步重点阐述 PFC 在三个方向上对于 Transformer 模型的启发, 包括注意力机制启发、生物编码原理启发、多感觉融合功能启发等, 并分别阐述这些启发在可解释性、低能耗性等方面的优势. 接着, 第 3 节中阐述 PFC 启发 Transformer 等神经网络模型的相关讨论和未来发展方向. 最后, 在第 4 节中进行总结与归纳.

1 前额叶皮层的生物分析

从结构角度出发, PFC 在不同尺度上呈现出多种特性. 在细胞类型尺度上, PFC 包含锥体神经元、中间神经元、星形胶质细胞和少突胶质细胞等, 它们各自发挥着不同的功能, 有学者研究了 PFC 不同细胞类型在联想学习过程中的群体编码特性和时间稳定性, 并识别出刺激、反应和新关联的异质群体编码[1]; 在局部神经环路尺度上, PFC 中神经元相互连接实现信息精确处理, 其中就有学者通过对绒猴 PFC 进行广泛的示踪映射, 从而发现斑片状和弥散状的两种类型投射在皮质和纹状体中呈地形排列, 揭示了局部投射的柱状结构、不同的层状分布模式、与其他区域的紧密相互连接等 PFC 局部和长距离回路的重要原理[2]; 在多种脑区尺度上, PFC 又可以细分为与视觉信息处理和眼动有关的尾侧 PFC、涉及动作顺序和目标转换的内侧 PFC、

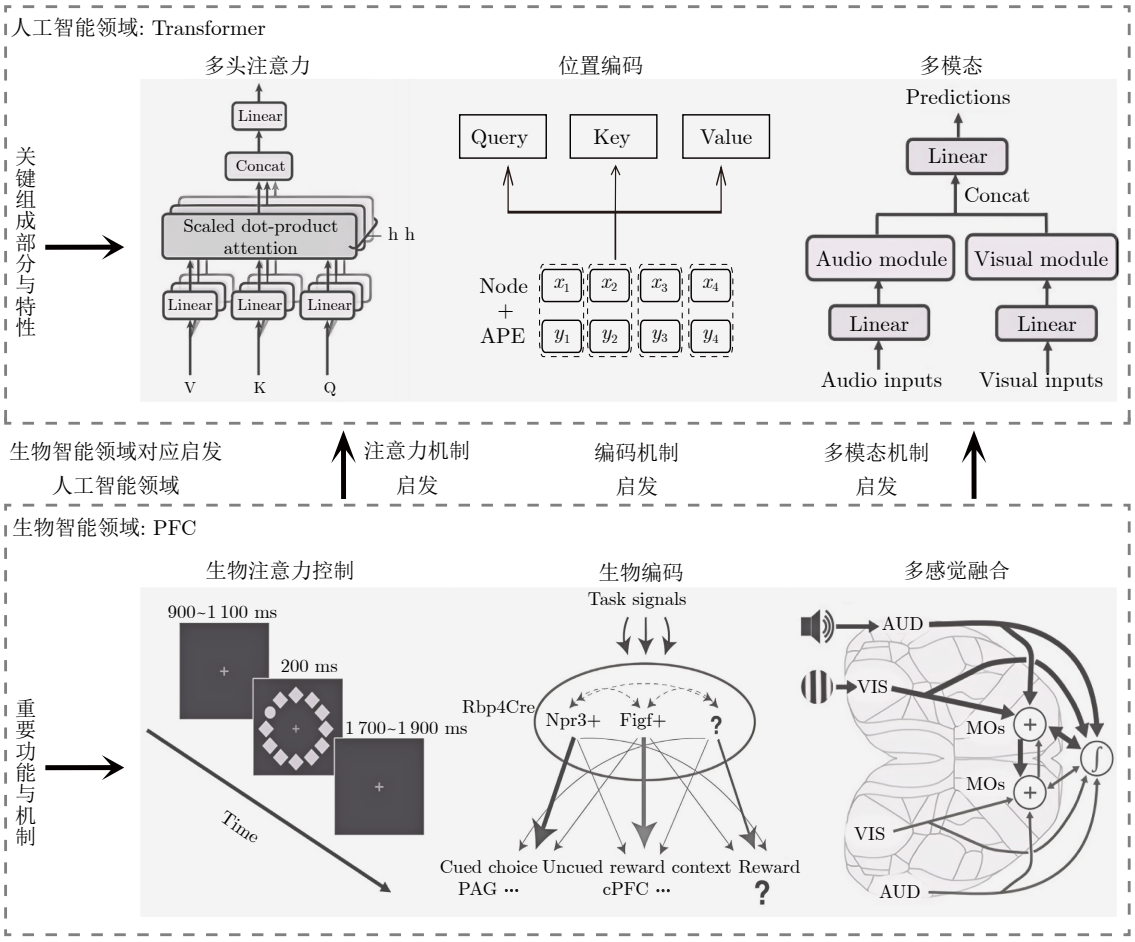


图 1 PFC 启发 Transformer 结构

Fig. 1 PFC-inspired transformer structure

与物体和声音关联学习有关的腹侧 PFC[3]; 在大范围脑区连接尺度上, 有学者通过研究发现 PFC 的一些脑区与抑郁症状相关, 如双侧背外侧 PFC、左侧背内侧 PFC 是“风险”区域, 右侧背内侧 PFC 是“韧性”区域, 并通过网络映射分析相关的功能和结构网络连接[4]. 正是因为这些结构层面上的特点, 才使得 PFC 拥有多种多样强大的功能.

从功能角度出发, 记忆、抽象思维、控制、决策都是人类非常重要的认知能力, 而这几种重要能力的产生都与前额叶皮层有着不可分割的联系, 以下将对 PFC 的这四种功能分别进行阐述.

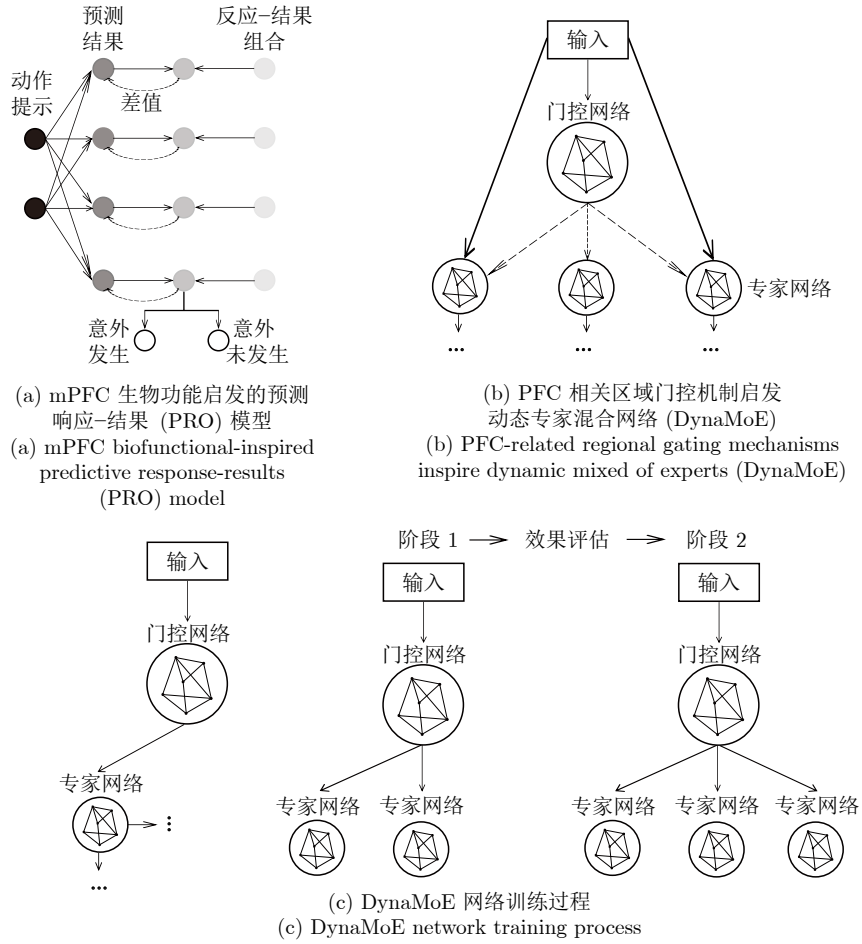
首先, 关于 PFC 在工作记忆中起到的关键作用, 有研究人员在实验中通过训练猕猴记住一个短暂出现的线索位置, 并在其消失 3 s 后根据记忆指示出消失线索的位置[5], 该项实验验证了 PFC 本身所具有的工作记忆功能. 研究结果表明 PFC 神经元在线索消失期间持续以较高的频率发射动作电位, 这种持续的神经活动在外部刺激消失后仍能有效控制行为、指导行为, 从而达到工作记忆的作用, 这种功能在人脑 PFC 中同样存在. 在抽象思维方面, PFC 也扮演着十分重要的角色. 相关研究表明, 人脑 PFC 中的单个神经元活动模式可以用来解码抽象的心理事件, 且 PFC 自身不同区域在处理抽象信息时也有着不同的功能分工[5], 这揭示出 PFC 在面对抽象信息输入时表现出的高度组织性和复杂性, 证明 PFC 确实拥有抽象思维的功能. PFC 在认知控制过程中同样有着十分重要的作用, 正如 Miller 等[6] 所指出: “认知控制源于在 PFC 中主动维持代表目标及其实现手段的活动模式. 它们向其他脑结构提供偏向信号, 其净效应是引导活动沿着神经通路流动, 建立输入、内部状态和输出之间的正确映射, 以执行给定任务”, 这表明 PFC 能够通过维护目标相关的神经活动模式和提供指引信号来管理和调节其他大脑区域, 以实现目标导向的行为和任务执行. PFC 及其相关网络中所蕴含的这种认知控制功能也将有望在未来的医疗领域中继续成为理解精神和神经疾病、开发新疗法的中流砥柱. 除此之外, PFC 的决策能力在组织行为过程中作为重要作用的分支点, 决定着具体行为将沿着哪一条路径发展. 已有研究中, 通过大鼠执行复杂的决策任务, 发现内侧前额叶皮层 (Medial prefrontal cortex, mPFC) 的背侧部分更多地参与处理有关主动决策的信息, 而腹侧部分更多地与动机因素相关[7], 这一结论进而验证了 PFC 确实存在决策功能.

正因为 PFC 拥有上述的这些重要功能, 所以 PFC 自然而然成为当下各领域的研究焦点, 例如目前 PFC 在类脑智能领域中就已经带来许多方面的应用, 这些应用对更好地理解大脑生物结构、功能与更好地启发人工智能领域发展有着重要意义. 有学者在传统的强化学习的基础上, 使用多巴胺系统训练大脑的 PFC 使其作为独立的学习系统, 并以此提出一个新的强化学习框架[8], PFC 在该框架的模型控制中发挥着重要作用. 还有学者在了解 mPFC 的功能及其在认知控制、行为预测中扮演的角色后受到启发, 提出一种新的 mPFC 功能理论模型, 即预测响应−结果 (Predicted response-outcome, PRO) 模型[9], 如图 2(a) 所示. 它以标准强化学习算法为核心, 以一个统一的理论框架解释 mPFC 已知的多种功能, 提出由生物启发人工智能模型并通过模型进一步阐释生物机制的重要思路. 还有研究发现 PFC 和相关区域可能通过采用类似门控的机制来控制大脑中的信息流, 并且以这种方式支持涉及多种模式的复杂行为[10−12]. 研究者基于相关发现, 提出一个名为动态专家混合 (Dynamic mixture of experts, DynaMoE) 的神经网络架构[13], 如图 2(b) 所示, 它将 PFC 中这种类似于门的机制、先前用于模拟 PFC 功能的循环神经网络 (Recurrent neural network, RNN) 和专家混合 (Mixture of experts, MoE) 架构、渐进式学习过程同步引入并结合, 从而实现不同模式的灵活终身学习, 网络训练过程如图 2(c) 所示.

以上列举的这些类脑智能方向的成果大多都是利用 PFC 包含认知控制在内的功能来启发类脑模型的研究, 这可以为神经网络模型增加至关重要的生物可解释性, 使模型中的神经元结构和功能更加接近真实生物层面的神经元, 从而为模型带来性能上的优化. 另外, 类脑智能的进步也会帮助人们更好地观察 PFC 等脑区的生物结构, 助力人们更深一步地理解大脑内部的生物相关机制, 形成一个正反馈的关系网, 整体逻辑流程如图 3 所示.

2 前额叶皮层启发 Transformer 模型

在上文中, 从结构、功能以及目前在类脑智能领域起到的重要作用入手, 对前额叶皮层进行了详细的介绍. 在人工智能领域, Transformer 是一个近年来非常热门且公认性能较好的模型, 它最初在自然语言处理 (Natural language processing, NLP) 领域中得到广泛应用, 并在近年来随着模型和技术的不断发展被拓展应用到计算机视觉、音频处理、强化学习等多个领域. PFC 和 Transformer 在各自的领域已经展现出十分重要的作用, 考虑生物大脑也是通过将皮层的皮质柱结构重复堆叠使用, 实现



(a) mPFC 生物功能启发的预测响应–结果 (PRO) 模型
(a) mPFC biofunctional-inspired predictive response-results (PRO) model

(b) PFC 相关区域门控机制启发动态专家混合网络 (DynaMoE)
(b) PFC-related regional gating mechanisms inspire dynamic mixed of experts (DynaMoE)

(c) DynaMoE 网络训练过程
(c) DynaMoE network training process

图 2 PFC 生物功能启发生物功能模型与神经网络架构

Fig. 2 PFC biofunctional-inspired biofunctional model with neural network architecture

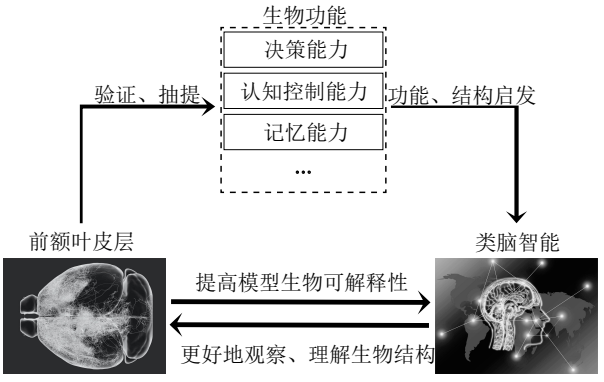


图 3 PFC 与类脑智能相互促进、共同进步

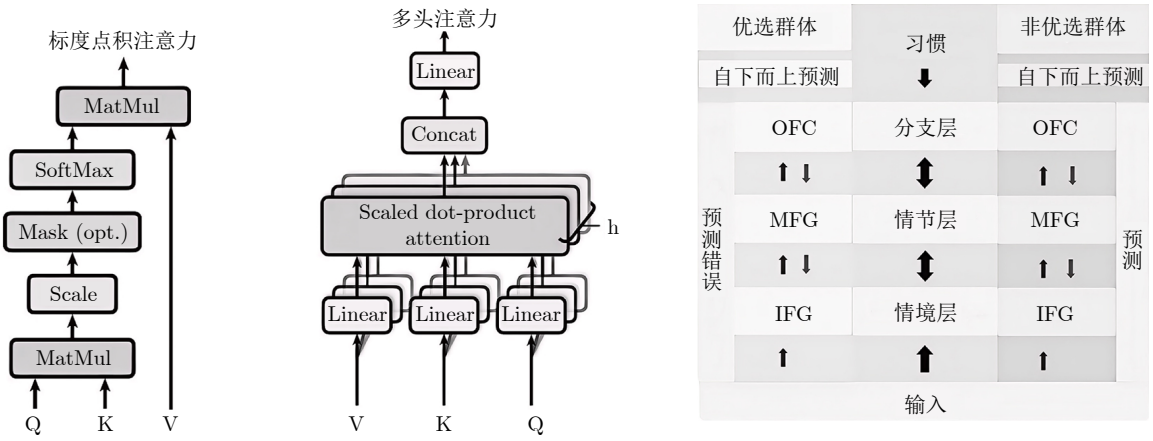
Fig. 3 PFC and brain-like intelligence promote each other and progress together

如视觉皮层、听觉皮层、运动皮层等多种高级认知能力, 这与堆叠 Transformer 实现通用大模型有着异曲同工之妙. 那么直观的想法是, 可否让二者进行“强强联合”, 通过 PFC 自身的功能与优势启发 Transformer 模型, 进而进一步提高模型的性能? 本节将对此进行介绍与分析.

下文将分别从注意力机制、生物编码原理、多感觉融合功能三个角度阐述 PFC 原理, 其可以分别对应启发 Transformer 模型在自注意力机制、位置编码、多模态三个方面的具体内容.

2.1 注意力机制启发

Transformer 的核心在于自注意力机制, 在自注意力机制中, 模型可以同时考虑输入序列中的所有位置信息, 并且可以根据每个位置的重要性不同进行相应的加权计算, 以便更好地捕捉序列中不同位置之间的依赖关系. Transformer 的自注意力机制包含标度点积注意力和多头注意力两部分[14], 如图 4(a) 所示, 它们可以使模型在每个序列位置上根据其他位置的信息来动态地调整权重, 进而做到并行地关注序列中不同位置之间的关系. 通过这样的方式, 模型能够同时实现处理长距离依赖和捕捉全局上下文的能力, 而无需像一些传统的神经网络那样逐步处理序列, 提升模型的表示能力和学习效果.



(a) Transformer 自注意力机制由标度点积注意力和多头注意力组成
(a) Transformer self-attention mechanism consisting of scaled dot-product attention and multi-head attention

(b) 基于 PFC 的分层注意力控制模型
(b) PFC-based hierarchical attention control model

图 4　PFC 与 Transformer 注意力相关模型架构

Fig. 4　PFC and transformer attention-related model architecture

近年来, Transformer 和注意力机制在自然语言处理、计算机视觉等领域取得成功, 有学者在这些基础上通过研究推动它们在生物信息学基因组数据分析领域中的应用[15]. 还有学者在自动驾驶的场景下, 基于 Transformer 注意力机制的 3D 点云运动物体分割方法, 通过设计双流网络融合点云和图像, 从点云和图像中提取多尺度特征信息, 进而提高运动物体的分割完整性[16].

随着生物领域的观察仪器和分析水平的不断发展, 目前已经能够确定 PFC 中同样存在着注意力机制. 例如, 学者们发现 PFC 在非空间注意力中起着关键作用, 其兴奋性状态可能受慢振荡动力学调节, 在实验过程中展示出非空间注意力根据 PFC 兴奋性状态的相位进行选择性调节, 这一过程通过影响感觉区域对相关特征的处理, 进而影响目标导向行为[17]. 在 PFC 的多种注意力中, 当前关注较多的是 PFC 中基于特征的注意力机制 (Feature-based attention, FBA), 它表现在当处于混乱的场景中时, PFC 可以通过基于特征的注意力来快速定位目标物体. 为更好地了解这种注意力机制具体是如何被用于寻找和选择目标的, 相关研究中对猴脑 PFC 中的腹侧斜纹前中回 (Ventral prearcuate, VPA) 区域展开测试[18]. 研究结果显示, VPA 中的神经元在整个过程中起到非常重要的作用, 它通过计算具有目标特征的物体位置信息, 并继续将这些信息发送给前额叶眼区 (Frontal eye fields, FEF), 从而引导眼球运动到相关刺激的位置. 这一研究结果对应于人类 PFC 中的前额叶下联合区 (Inferior frontal junction, IFJ), 因此也同样适用于人类 PFC. 除此之外, 还有研究人员针对基于特征的注意力机制在人类 PFC 类别诱导的全局效应展开研究, 发现这种全局效应主要由 PFC 中的 IFJ 区域提供反馈调控实现[19]. 这一发现意味着 PFC 的功能不仅仅局限于低级的特征相似性处理, 而且还应该包括较为高级的类别处理, 这种功能有助于在复杂场景中根据目标类别的不同灵活地调节注意力. 以上提出的两项研究均是围绕着 PFC 基于特征的注意力机制展开的, 当然, PFC 所拥有的注意力机制并非仅仅体现在特征捕捉中, 在跨模态注意、情绪调节乃至社会认知等方面也会有所体现, 通过这些相关研究的研究结果可以验证 PFC 中注意力机制是切实存在的.

现如今, 越来越多的研究人员开始对人类注意力和 Transformer 自注意力机制进行比较分析, 其中包括选择性注意力、语境理解等相似点, 以及容量限制、注意力路径、意向性等不同点[20]. 那么在了解 PFC 和 Transformer 各自所拥有的注意力机制后, 我们就是否可以通过 PFC 在生物层面的注意力机制启发 Transformer 的自注意力机制展开研究. 已有研究表明, 受 Transformer 注意力机制的启发, 采用由功能特定的“注意力头”执行新兴句法计算, 可以差异性地预测大脑中包含 PFC 在内的特定皮层区域的活动[21]. 结果表明, 语言模型和皮层语言网络在处理自然语言时可能会收敛于相似的特定功能. 此外还有学者在了解大脑中“自注意力”可以通过状态空间模型高效、高速率地实现后, 发现基底神经节、小脑、PFC 等大脑结构与 Transformer 中自注意力机制的结构相结合可以实现复杂的注意力和动作协调功能[22]. 还有生物研究者给出一种基于 PFC 的分层注意力控制模型. 这个模

型的算法原理基于信息理论和认知控制理论, 认为注意力控制可以被概念化为由层级有序的控制过程组成, 每个过程负责基于不同时间尺度的信息选择行动[23], 如图 4(b) 所示. 与此同时, 近年来人工智能领域的研究人员也对分层注意力产生浓厚的兴趣并取得进展. 其中就诞生出名为 BViT (Broad attention based vision Transformer) 的视觉 Transformer 模型[24], 它利用广泛注意力机制, 通过在不同层间整合和利用注意力关系来提取有用信息进而改进性能, 其核心的广泛注意力模块在结构层面上利用全局连接促进层与层之间信息的传递和整合, 并增加不同层之间的路径连接来获取更丰富的信息. 无论是 PFC 中的分层注意力控制模型, 还是视觉 Transformer 模型 BViT, 二者都是围绕着分层注意力这个算法原理本质展开的, 这可以为我们提供值得进一步探索的方向, 即围绕着分层注意力将二者进行结合, 使 PFC 在处理和分配注意力的角度上为 Transformer 的模型架构提供灵感.

除此之外, 还有研究者在了解大脑中相关机制与注意力之间不可分割的关系后得到灵感, 从架构和算法原理层面出发, 提出一种名为多模态通道式注意力 Transformer (Multimodal channel-wise attention Transformer, MCAT) 的模型[25]. 后续测试中更是验证了 MCAT 中加入的注意力模块对提高模型融合神经网络的性能是必要的. 还有研究人员受 PFC 相关的执行注意理论启发, 假设 Transformer 中的自注意力机制可能是其工作记忆容量受限的原因, 并通过后续训练模型进行 N-back 任务, 为 Transformer 自注意力机制与工作记忆容量的关系提供见解[26]. 还有研究者设计出一个开放的视觉皮层–基底神经节–前额叶皮层环路介导的自上而下视觉注意控制模型[27], 这个模型通过基底神经节与 PFC 的交互实现更加动态和自适应的注意力控制架构, 选择由目标和干扰之间的信噪比来引导注意力机制而并不只是简单地将注意力直接指向目标特征. 这种注意力动态调整的思路可以很好地启发 Transformer 在模型架构上的进步, 使之有望形成类似于实时注意力微调的功能, 提高模型在任务执行过程中的效率.

总结来说, 在上面所给出的研究案例中, 研究人员对人类注意力和 Transformer 自注意力机制进行比较, 发现相似点与不同点. 在此基础上, 一方面, 受 Transformer 注意力机制启发, 有研究可预测大脑特定皮层区域活动, 且大脑结构与 Transformer 自注意力机制结合能实现复杂功能. 另一方面, 生物领域基于 PFC 的分层注意力控制模型与人工智能领域的 BViT 视觉 Transformer 模型均围绕分层注意力展开, 为 PFC 与 Transformer 的结合提供方向, 体现从不同角度探索两者结合的逻辑, 有望使 PFC 在 Transformer 模型架构的注意力处理和分配方面发挥启发作用.

在本节中, 我们分别从 PFC 的生物注意力机制和 Transformer 的自注意力机制角度, 盘点了在注意力机制方面 PFC 可以为 Transformer 在架构和算法层面带来的启发. 无论是已有的研究进展还是未来前景启发, 它们的共同之处都在于通过面对具体的情境或者目标时, PFC 展现的注意力控制能力可以启发 Transformer 模型根据特定情境、目标输入更好地分配注意力, 从而提高模型快速、准确、灵活地从复杂输入中提取有用信息的能力. 同时, Transformer 核心的自注意力机制也可以启发 PFC 中的生物注意力机制, 从而构建出更加精确的大脑模型并揭示大脑在处理注意力时的动态变化, 帮助我们更好地理解复杂认知过程中的神经基础, 整体逻辑关系如图 5 所示.

2.2 PFC 生物编码原理启发 Transformer 位置编码

传统的 Transformer 模型大多应用于自然语言处理领域, 但是其核心的自注意力机制是不能获取一句话中词语位置信息的. 众所周知, 对一句话中出现的同一个词, 如果词语出现位置不同, 那么含义可能会发生翻天覆地的变化. 因此提出的 Transformer 位置编码正是为给模型提供输入序列差异而诞生. 近年来, 随着对 Transformer 理解的不断加深, 位置编码的研究也有进一步的发展. Transformer 的位置编码可以分为绝对位置嵌入 (Absolute position embedding, APE) 和相对位置嵌入 (Relative positional embedding, RPE) 两类, 前者通过将不同位置的词语映射为不同的向量来表示它们的绝对位置, 后者利用当前词语与其他词语之间的位置关系来编码位置信息. 对于绝对位置编码, 它为输入序列中的每个位置分配一个唯一的编码向量, 使每个位置都有一个固定的、绝对的位置信息, 例如在最早的 Transformer 模型中研发者就是使用正弦和余弦函数完成这一操作. 目前, 绝对位置编码被更多地应用于 AI 技术或工具中, 如 GPT3[28–29]、OPT[30] 等. 对于相对位置编码, 有学者在对相对位置向量背景的研究过程中提出使用位置矩阵的方法来改进相对位置编码[31]. 还有学者将目前较为流行的几种位置编码方法, 如 Rotary、ALiBi、T5 相对偏差 (T5′s relative bias) [32] 等进行归纳总结. 其中 Rotary 方法的核心在于将位置信息用旋转矩阵

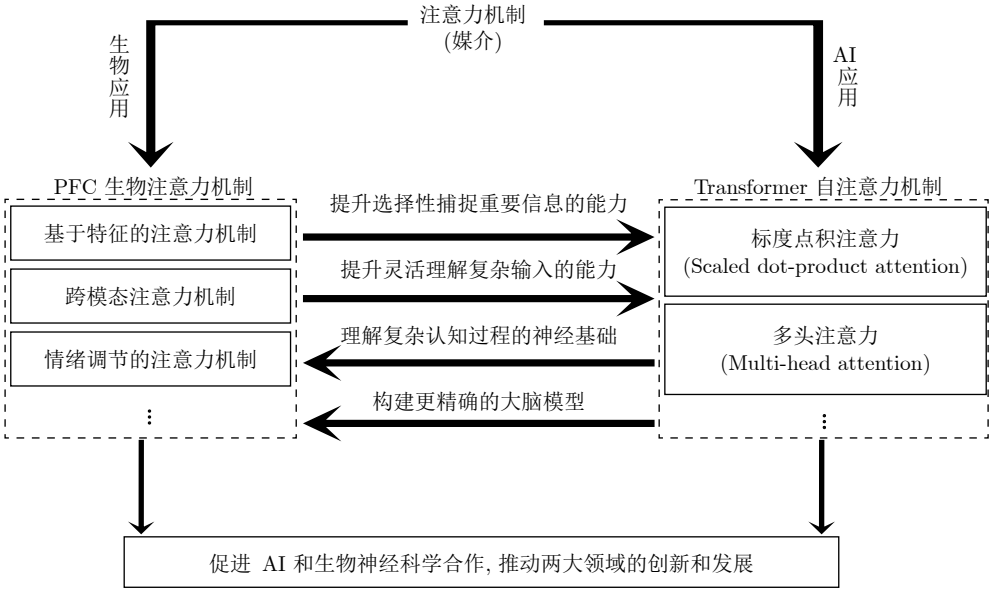


图 5　PFC 与 Transformer 以注意力机制为媒介相互启发

Fig. 5　PFC and transformer inspire each other through the medium of the attention mechanism

的形式嵌入 Query 和 Key 向量, 通过点积计算出注意力使其只取决于标记间的相对距离进而有效地实现相对位置编码, 目前 Rotary 方法主要在 PaLM[33] 和 LLaMA[34−35] 中得以应用. 除此之外, T5 相对偏差则是使用一个查找表将两个位置之间的相对距离映射为一个标量偏差值, 而后将偏差值添加至注意力分数的计算过程中. 研究者基于 T5 相对偏差方法, 将标量相对位置编码 (Scalar relative position encoding, SRPE) 与设计的自适应 T5 Transformer (Adaptive T5 (AT5) transformer) 模型相结合[36], 使用标量嵌入相对位置并使用固定的启发式算法对相对位置进行“桶化”, 使得同一桶中的位置能够共享相同的嵌入信息, SRPE 的结构如图 6 所示. 设计出的 AT5 Transformer 模型则克服原始 T5 Transformer 模型固定的分桶函数和独立学习的标量嵌入两大不足, 有效地控制欠拟合和过拟合风险. ALiBi 方法与 T5 相对偏差方法类似, 所不同的只是 ALiBi 从注意力分数中减去一个随着 Query 和 Key 之间距离线性增长的标量偏差, 从而能够更好地捕捉长程依赖关系, ALiBi 方法主要在 BLOOM 中得以广泛应用.

PFC 作为大脑皮层中负责复杂认知功能的区域, 它的生物编码机制在这些功能的产生中起着至关重要的作用. PFC 的生物编码主要体现在它的神经元以特定方式编码和处理信息, PFC 可以通过神经元放电、神经递质调节、突触可塑性等多种方式完成简单或复杂的编码过程. 它的生物编码过程也与 Transformer 模型的位置编码在许多方面存在着

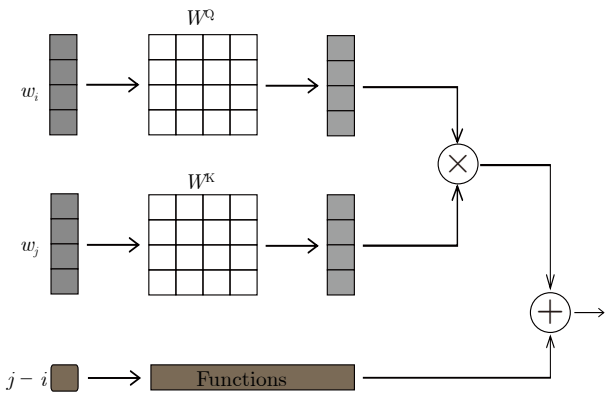


图 6　标量相对位置编码 (SRPE) 原理

Fig. 6　Principle of scalar relative position encoding (SRPE)

相似之处, 因此如何将二者结合带来启发成为一个值得探究的问题, 目前对这个问题的解答也诞生出一些研究成果与可行思路.

首先, 一部分研究者从 PFC 的灵活性和对输入信息的辨识能力方面入手. 有研究者通过记录两只实验状态猕猴的 PFC 中单个神经元的活动, 得出 PFC 中神经元不仅可以从抽象的行为中提取出编码规则, 而且还能够灵活地在不同的编码规则之间进行灵活切换的结论[37]. 这一发现启发我们可以利用 PFC 的这种灵活性与 Transformer 位置编码的算法原理进行结合, 从而提高模型理解序列中元素位置关系的能力. 另外, 已有工作中得到的神经编码模型在两个大脑区域 (视觉皮层 V4 和前额叶皮层 IT) 中效果不错, 但在大脑其他部分的大脑编

码方面仍需要大量的努力来改进. 在此背景下, 通过参考在前额叶皮层表现不错的神经编码模型, 研究者对基于图像和多模态 Transformer 是否可以准确地对整个大脑进行 fMRI (Functional magnetic resonance imaging) 编码进行研究[38]. 还有研究者提出 PFC 快速编码新功能联系的能力对于适应性行为、工作记忆至关重要, 这往往需要打破特征之间的历史绑定用以编码新特征, 且突触强度、神经元调谐等可能会发生根本性变化[39]. 基于此, 我们可以为 Transformer 的模型架构加入类似于 PFC 的灵活调谐能力, 使位置编码更具动态性, 以面对不同输入时可以更加灵活地处理长短期特征, 达到更有效地整合不同时间尺度信息的目的. 此外, 通过借鉴 PFC 编码在时间和位置信息上的灵活性, 还有研究者为提出的 AT5 Transformer 模型在算法原理层面上设计一个更加动态和可学习的分桶函数, 使模型可以自动调整以适应不同的任务[36].

其次, 部分研究者试图利用 PFC 的生物编码提高 Transformer 模型在处理多种实际任务时对任务上下文的理解能力和特征提取能力. 模糊神经网络 (Fuzzy neural network, FNN) 虽然因为可解释性和自学习能力而备受关注, 但是其在解释高维非结构化数据问题上遇到困难. 为解决这一问题, 研究者提出一种基于变分自编码器 (Variational autoencoder, VAE) 和 FNN 的可解释图像分类模型——VAE-FNN[40], 输入经过模糊神经网络分类器 (Fuzzy neural network classifier, FNNC) 得到分类结果, 该分类器基于从视觉皮层提取的特征用来模拟顶叶和前额叶皮层的推理分类过程, 模型中使用的 VAE 还可以根据实际应用场景, 使用视觉 Transformer (Vision transformer, ViT)、编码器–生成对抗网络 (Generative adversarial network, GAN) 生成器等进行替代. 除此之外, 还有研究者发现 PFC 拥有准确编码任务语境信息的能力并能够保持高保真度[41], 因此可以进一步启发我们将 PFC 的语境编码能力与 Transformer 的整体架构结合来增强模型对任务上下文的理解, 从而更好地对给定的输入进行处理. 相关研究中还提出一种新的自监督学习信号——位置标签, 通过位置标签作为自监督学习信号来训练视觉 Transformer, 可以识别输入图像中各个编码片段的位置, 实现一种有意义的自监督任务[42]. PFC 中神经元对特定的空间位置会产生对应的特定反应, 这与提出的位置标签的概念有相似之处, 可以在后续研究中考虑将其整合至 Transformer 模型的原理中. 另外, 额顶叶皮层在预测编码机制中的作用启发基于 Transformer 的语言模型. 研究表明当语言模型被优化为预测附近的单词时, 它们的表现仍然逊色于人类大脑的语言处理能力. 预测编码理论提出, 其原因在于人类大脑通过跨越多个时间尺度和层次的预测来处理语言信息. 额顶叶皮层特别负责预测更高层次、更长时间范围和更多上下文相关的表征. 这些多尺度的神经科学发现, 可以用来改进大语言模型[43], 而前额叶皮层和额顶叶皮层在语言处理和预测编码机制中有着密切的联系和共同作用, 因此也可以有类似的应用.

最后, 还可以通过时序信息嵌入来将 PFC 生物编码和 Transformer 位置编码建立联系. 例如在相关文献中, 提到外侧前额叶皮层 (Lateral prefrontal cortex, lPFC) 中的神经元激活序列 (Neuronal activation sequences, NAS) 在时空复杂任务中的工作记忆 (Working memory, WM) 编码机制[44], 那么这里 lPFC 中的 NAS 在时空复杂任务中体现出的时序信息引起更多的关注, 是否可以将其应用至 Transformer 位置编码中以实现时序信息的嵌入, 使模型能够更好地处理时空复杂的输入数据呢?对此, 我们对与时空信息相关的 Transformer 模型进行研究并收集到现有的研究进展. 有研究者提出一种新的位置编码算法——基于离散傅里叶变换 (Discrete Fourier transform, DFT) 编码, 它的每个位置编码与相应的位置函数都是一一对应的, 因此也称为“忠实编码”. 基于“忠实编码”, 文献 [45] 进而提出一个名为 DFStrans 的模型, 它结合 1D 多头卷积神经网络和类似 Transformer 的时空结构, 通过对时空依赖性进行建模来提供局部和全局的诊断评分. 还有研究者提出一个名为 STTRE 的模型[46], 它通过重新结构化多头注意力机制, 从而能够更好地利用相对位置编码, 增强对时间序列数据中隐含时空依赖性的捕捉能力. 更有研究者在视频处理方面提出一个基于相对位置嵌入的用于动作识别的空间和时间解耦的 Transformer 模型——RPE-STDT[47], 它通过两步解耦过程使模型能够有效地捕获空间和时间依赖性, 从而获得更全面和信息量更大的视频表示. 以上研究中提出的各种 Transformer 模型都与时空信息密切相关, 通过与 PFC 在时空复杂任务中时序信息相结合, 相信可以使模型在处理多种实际任务时更好地捕捉给定输入中的时序信息与空间信息, 并优化模型性能.

总结来说, 在上述的案例中, 研究人员从多方面探索 PFC 的生物编码与 Transformer 的位置编码结合. 一方面, 着眼于 PFC 的灵活性和输入信息辨识能力, 启发与 Transformer 位置编码结合以提

升位置关系理解等能力, 并设计动态分桶函数. 另一方面, 利用 PFC 生物编码提高 Transformer 对任务上下文的理解和特征提取能力, 如在模型中替代部分模块、结合语境编码能力、整合类似概念至 Transformer 等. 最后, 通过时序信息嵌入建立联系, 研究与时空信息相关的 Transformer 模型, 并结合 PFC 时序信息从而优化性能. 这些研究体现了从不同角度探索两者结合的逻辑, 为 Transformer 模型发展提供新方向.

本节中, 首先对 Transformer 的位置编码进行详细的介绍并盘点目前较为流行的一些位置编码实例. 之后从 PFC 的灵活性与辨识能力、PFC 语境信息编码能力、PFC 时空信息处理能力三个角度阐述 PFC、Transformer 在各自领域中的研究发现, 并且从编码的算法原理角度对 PFC 启发 Transformer 的研究进展与可行思路进行讨论, 整体框架图如图 7 所示. 这些成果的创新点都围绕着提高 Transformer 对输入信息的理解、处理等能力来进行深入挖掘. 不足之处则在于还需要更多的实验结果来对给出的思路进行验证.

2.3　PFC 多感觉融合功能启发 Transformer 模型

PFC 作为大脑中拥有认知控制、任务决策等功能的重要区域, 它在实现这些功能前会从外界接收多种输入感知信息, 这些感知信息可能来自于生物的视觉、听觉、触觉等多种途径, 并且相对复杂、容易串扰, 因此将这些输入信息进行合理整合就显得尤为重要. 本节中将会提到的 PFC 中拥有的多感官融合能力正是解决相关问题的一条路径, 它可以指导大脑将不同来源的信息进行整合, 以形成一致、连贯的感知, 进一步启发 Transformer 的多模态信息融合能力.

目前, 在对 PFC 的多感官融合功能的研究工作中诞生了许多新的成果. 有研究者在研究 PFC 的多模态功能的过程中得出 PFC 中的运动皮层部分 (Motor cortex, MOs) 是整合由多感官功能得到的多模态表现的关键皮质区域这一结论, 并提出一个通过皮质启发的视听整合通路模型[48]. 在这个通路中, 视觉和听觉的单一感觉信息通过视觉和听觉感觉皮质传递至 MOs, MOs 中神经元将来自视觉和听觉的刺激通过“加法”方式进行整合, 进而引导下游回路通过“积分”做出具体决策. 还有研究者着眼于注意力如何通过在皮层不同层次上运作的两种不同机制——注意力前机制、刺激后机制来控制多感官感知[49], 研究发现 PFC 的作用主要体现在刺激后机制中, 它可以利用贝叶斯建模的相关方法将视觉和听觉信号整合为空间估计, 并且可以在感觉的整合与分离仲裁中发挥关键作用[50]. 这一发现说明 PFC 可以动态地整合来源于不同感官的信号, 进而在感知推断过程中起到重要的调控作用. 除此之外, 还有在虚拟多感官环境中进行的相关研究, 得出额叶前端和额叶下端在多感官融合过程中起到重要作用的结论[51]. 如一方面, 实验中发现当给到前运动皮层的振动触觉刺激接近阈值时, PFC 中的前运动皮层会显著参与处理过程并对微小刺激的主观感知造成影响. 另一方面, 还发现当感觉一致性线索与情境预期冲突时, 额叶下端会参与调解这种冲突并调整信息整合方式[52−53]. 在记录 PFC 中靠近弓状沟区域的单元活动时, 发现 PFC 中前额叶眼区的单个神经元在实验过程中表现出很大的异质性, 使用分解主成分分析 (De-mixed principle component analysis, dPCA) 更是揭示出神经元群体中存在着稳健的刺激模态信号和选择信号[54]. 这表明, 在知觉决策任务中起着关键作用的 FEF 同样可以区分感官输入类型, 在跨感官整合功能方面也有着不错的建树.

在人工智能领域, 多模态模型不仅提升单一模

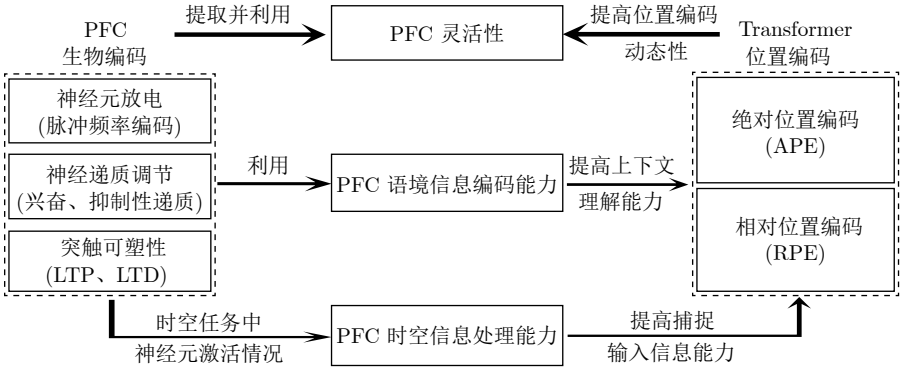


图 7　PFC 生物编码过程启发 Transformer 位置编码

Fig. 7　PFC biological coding process inspired transformer position coding

态模型在具体任务中的效果, 而且还推动人工智能技术的发展, 使其逐步向着人类多感官信息处理的方式接近. 通过开发具有理解、推理、学习等能力的计算机智能体, 使之可以像人类一样使用多感官模态感知来与世界进行互动, 一直是人工智能领域的宏伟目标[55]. 此外, 神经科学和心理学的现代进展表示, 多感官输入对认知功能[56] 至关重要, 因此建模和理解多模态相互作用可以为开发受人类大脑启发的更高智力水平智能体开辟道路[57]. 多模态 Transformer 模型正是实现这一宏伟目标的关键, 它是一种能够处理和融合包含图像、文本、音频在内的多种不同模态数据的深度学习模型. 基于经典的 Transformer 模型, 更是在多模态能力的加入后能够从多角度理解和处理信息, 使模型在复杂任务中能够利用多种信息源的优势来弥补单一模态的局限性. 目前, 基于 Transformer 模型的多模态交互仍然是一个深受关注的研究方向, 以自注意力机制为核心的 Transformer 模型除可以捕获比其他神经网络模型更多的全局信息, 还可以较为自然地捕获多模态数据之间的交互过程[58]. 在这样的背景下, 许多多模态 Transformer 模型被陆续研发出来, 例如有研究者提出一种基于模态翻译的 Transformer 模型[59], 有效地实现了模态间的端到端融合, 还有研究者提出一种分层交互多模态 Transformer 模型——HIMT[60], 通过目标检测方法从图像中提取出具有语义概念的显著特征. 在对多模态情感分析领域的研究中, 还诞生一个名为 TMBL 的模型[61]. 它将模态特征分类为模态不变特征和模态特定特征两种类型, 并通过设计出的多头绑定转置注意 (Multi-head binding transposed attention, MBTA) 和门控绑定前馈网络 (Gated binding feed-forward network, GBFN) 加入 Transformer 模型中得到了优化的 All-in-One Transformer 架构, 进而能够更好地捕获交互模态之间的特征. 在 CMU-MOSEI 数据集上得到的实验结果验证了 TMBL 模型相较于过往模型在许多指标上更具优势, 显示了其充分提取模态特征的能力.

在盘点 PFC 多感官融合功能与多模态 Transformer 模型在各自领域中的发展现状后, 可以说二者不仅在学术层面具有很大的研究价值, 而且还对我们人类的日常生活起到十分重要的作用. 如此有意义的两个领域若是能够相互结合、相互启发势必会推动它们更进一步发展, 目前在二者结合的研究中也确实诞生了许多新成果和新思路, 下面将对此进行总结.

有研究者借鉴大脑中包括 PFC 在内的区域中的多感觉整合机制, 以此改进并优化现有的多模态融合模型, 提出一种名为多模态通道式注意力 Transformer (MCAT) 模型[25]. 还有研究者在研究中遇到因为不同模态序列接收频率不同所导致的多模态流异步问题, 针对此问题他们借鉴人类感知系统中优越的多模态融合过程, 通过研究包含 PFC 在内不同器官的感觉信号通过感觉神经元、存储神经元的统一处理和整合过程, 进而提出一种针对异步多模态序列的多模态融合和源模态协同强化 (Modality co-reinforcement, MCR) 的新方法[62]. 除此之外, 还有研究者通过在空间、频谱和时间域上应用多头注意力进行不同模态 (如, MI (Motor imagery), SI (Speech imagery), VI (Visual imagery)) 下心理意象的 EEG 多重分类, 从而提出一种用于不同模态下 EEG 分类的多尺度卷积 Transformer 模型[63], 如图 8 所示. 结果显示在 PFC 和补充运动区 (Supplementary motor area, SMA) 中存在共同的神经活动, 这表明这些区域在不同类型的心理意象中都被激活, 其中 PFC 在多模态任务中起到重要作用. 将 PFC 多感官融合功能和多模态 Transformer 模型的架构和算法结合应用至情感识别领域也是当前热门的研究方向. 有研究者提出一个以生理信号预处理、辅助模态特征优化、主模态特征增强、情绪预测四个步骤为核心的基于生理信号的多通道情绪识别框架[64]. 在整个框架中应用一种改进的交叉模式 Transformer (Modified cross modal transformer, MCMT), 其中交叉模式是在研究 PFC 对情绪激活活动的关联性过程中发现的, 因此在使用辅助模态时注意力权重将更集中于这些情绪相关的脑区, 进而使 PFC 在多模态框架中起到显著提升情绪识别性能的作用. 还有研究者发现多脑区的信息融合可以模拟大脑不同区域间的交

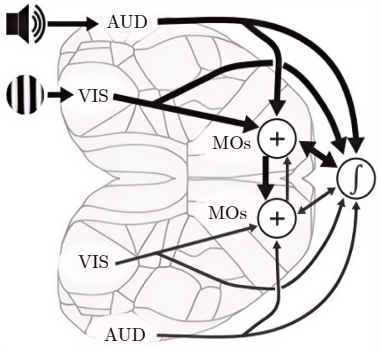


图 8 不同模态下 EEG 分类的多尺度卷积 Transformer 模型

Fig. 8 Multi-scale convolutional transformer model for EEG classification in different modalities

互, 这与 PFC 在情感和认知任务中整合信息的功能相似, 因此在多脑区信息融合的启发下提出一种新的模型, 旨在通过融合多模态数据来进行情感识别[65]. 以上总结出的这些成果展现了 PFC 自身和它的多感官融合功能对多模态 Transformer 在注意力、异步问题、分类问题、情感识别等多方面的启发与帮助, 更说明了 PFC 启发多模态 Transformer 的广泛性及可行性.

还有许多现有的研究成果为 PFC 多感官融合与多模态 Transformer 的算法结合提供不错的思路. 有研究者在研究中发现前额叶−丘脑 (mPFC-Re) 回路在处理多感官和内部线索整合上表现出色, 可用于区分不同的社会刺激[66]. 因此将这种回路的整合机制引入 Transformer 的模型架构中将有很大潜力提高模型在不同任务中处理文本、图像、音频、视频等多模态数据时的表现. 对于多模态 Transformer 在图像字幕领域的应用, 有研究者提出一种新的用于图像字幕的多模态变换 (Multimodal transformer, MT) 框架, 通过适当地深度堆叠注意块形成一个深度编码器−解码器模型, 同时捕获每个模态内的自注意和不同模态间的公共注意, 从而更好地理解视觉对象之间的复杂关系[67]. 图像字幕本身着眼于视觉和文本信息, 因此利用 PFC 对视觉等多感官信息进行整合可以进一步启发 MT 框架. 此外, 生物研究者发现 PFC 中的运动皮层部分 (MOs) 是整合多模态表现的关键皮质区域, MOs 的响应通过加性整合视觉和听觉信号形成[48]. 在相关文章中还提出注意力通过在皮层不同层次上运作的两种不同机制控制多感官感知同样涉及到了注意力的动态调整[49], 这与其他文章中提出的 TMBL 模型中的多头绑定转置注意力模块 (MBTA)[61] 可以在结构层面进行结合, 进而促进多模态 Transformer 模型性能的提升.

总结来说, 上述提到的这些成果表明 PFC 自身及其多感觉融合功能对多模态 Transformer 在注意力、异步问题、分类问题、情感识别等多方面有启发和帮助, 许多现有研究成果为 PFC 多感觉融合与多模态 Transformer 的算法结合提供思路, 如将前额叶−丘脑回路整合机制引入 Transformer 架构、利用 PFC 整合视觉等多感觉信息启发图像字幕的 MT 框架、将注意力动态调整与 TMBL 模型的 MBTA 结合等, 都说明了 PFC 的多感觉融合功能启发多模态 Transformer 的广泛性及可行性.

本节中, 我们首先盘点了 PFC 多感觉融合功能与多模态 Transformer 模型在各自领域中的发展现状, 在了解二者的基本内容后, 我们随即对 PFC 多感觉融合在算法、模型结构等方面启发多模态 Transformer 以及二者结合推动其他应用领域进步进行阐释, 并给出一些目前已有的研究进展和一些值得继续探究的研究思路, 整体的逻辑结构图如图 9 所示.

3 讨论与展望

近年来, 神经科学和人工智能领域各自取得巨大的进步, 神经科学的研究更是激发许多关于大脑结构和功能的新思考, 并切实推动了类脑智能领域的不断进步[68]. 在前文中已经阐述可以根据 PFC 拥有的功能, 通过不同的途径来给予 Transformer 模型新的启发, 考虑到模型的复杂性与任务的复杂性要对齐才能最大化发挥其优势, 接下来本文将继续针对 PFC 启发 Transformer 的应用领域进行总结.

3.1 PFC 启发 Transformer 多领域应用

在前文中我们已经盘点了 PFC 可以与 Transformer 模型通过注意力机制、编码方式、多模态等途径进行结合, 使 PFC 在生物层面的功能可以启发人工智能领域 Transformer 模型的优化与进步, 那么本节中将盘点 PFC 具体可以启发 Transformer 在哪些领域进行应用.

1) PFC 可以为 Transformer 在自然语言处理 (NLP) 领域的应用提供思路. 自然语言处理是指计算机处理和理解人类自然语言的能力, 其中涉及到文本分类、情感分析等多方面的内容. 目前, NLP 仍然深受大众关注并在不断地拓展创新中, 但是随着研究的不断深入, 包含 NLP 与语言神经科学之间缺乏交流[69] 在内的问题也不断地浮现出来. 虽然 NLP 的研究目的在于提高模型在各种标准化基准上的性能, 但只有少数研究中考虑神经认知研究中给出的见解[70−71]. 因此, 将在生物领域神经认知研究中得出的结论与 NLP 研究融合起来, 很可能会成为 NLP 更进一步发展的关键. 例如有研究者建立两个基于 BERT 的深度语言模型 (Deep language model, DLM), 通过使用下一个句子预测 (Next sentence prediction, NSP) 任务来建立语言理解, 进而可以探讨话语层次理解的神经机制. 这展示了 NSP 预训练能够增强模型与真实大脑数据的对齐, 阐释了非经典语言区域为高层次语言理解带来的贡献, 例如不一致文本整合会被认为涉及双侧背外侧 PFC 网络等[72]. 此外, 情感分析作为一种常见的 NLP 任务, 它在人工智能领域发挥着重要作用[73]. 例如处于认知神经科学和 NLP 交叉研究领域的认知启发式 NLP 任务, 近年来作为一种新颖的多模

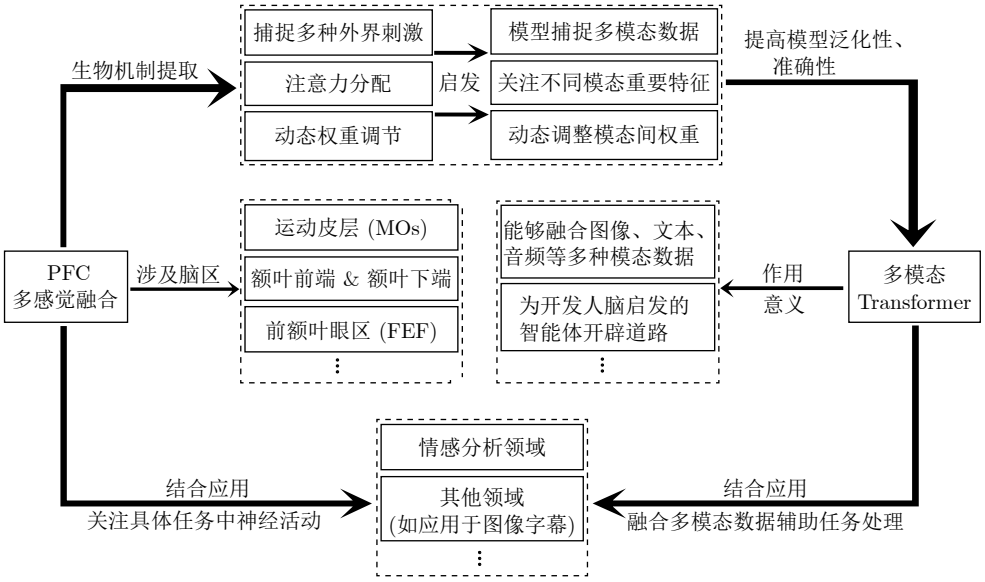


图 9 PFC 多感觉融合与多模态 Transformer 逻辑结构图

Fig. 9 PFC multisensory fusion with multimodal transformer logic structure diagram

态方法在情感分析领域取得优异的成绩[74]. 有研究者在相关研究中提出一种用于提高受认知启发的情感分析性能的多模态框架 CogAware[75], 它利用 RoBERTa 中的 Transformer 架构获取更高质量的词嵌入, 并提升情感分析任务的准确性和鲁棒性. 在前文中, 我们已经总结了 PFC 可以启发 Transformer 在情感分析领域不断进步, 因此未来可以进一步尝试通过 PFC 的多感官融合功能启发这里提出的 CogBERTa 模型中的 Transformer 结构部分, 使模型在情感分析领域更准确、快速. 此外, NLP 任务中还包含讽刺识别、抑郁症分类、文本毒性检测等多种子任务[76], 这些子任务也拥有着与 PFC 相结合的可能与前景[77]. 研究人员可以比较赞美与讽刺两种不同意图条件下的大脑活动和连接模式, 并通过频谱分析, 得到在讽刺表达时 PFC 等区域呈现出与积极情感表达不同的活动模式的结论, 从而揭示出大脑活动模式与讽刺识别之间可能存在的关联[78]. 在抑郁症识别分类的研究中, 深度学习算法得到广泛应用, 通过分析大量包含 fMRI 在内的神经影像学数据, 来寻找抑郁症患者和健康人群之间的差异模式. PFC 作为大脑中与情感、认知等功能密切相关的区域, 它呈现出的特征变化可以为抑郁症的识别分类提供重要依据[79]. 在文本毒性检测领域中, 研究者提出一种多语言毒性文本分类器, 采用不同损失函数和多个预训练模型的策略, 解决多语言环境下毒性评论分析的三个主要挑战: 多语言特性、标注数据缺乏和样本分布不均衡. 相信未来可以进一步研究注意力机制在 PFC 中的运作方式, 以及它如何影响对文本毒性的感知[80].

2) PFC 还可以与 Transformer 在视觉处理领域进行结合. 视觉 Transformer (ViT) 是一种 Transformer 架构的深度学习模型, 它可以直接处理图像的像素值, 而不需要进行特征提取和降维. ViT 的出现标志着图像处理, 尤其是神经成像领域的重大转变[81]. 这一创新模型与传统的卷积神经网络不同, 它结合最初为自然语言处理而设计的机制, 如自注意力机制, 从而实现更细致、更全面的图像分析[82]. ViT 的架构通过多个转换层将图像分割为多个片段进行处理, 从而实现与图像片段的空间位置无关的深入分析. 一方面, 视觉处理领域的 Transformer 模型原理可以使我们更好地了解 PFC, 例如有研究者基于不断发展的 ViT 模型和传统的生成对抗网络开发出一个新的条件生成对抗网络模型——cEViT-GAN, 它使用高效的 ViT 模型作为生成器和判别器, 并选择分块式自我注意层以捕捉人脑结构和功能磁共振成像 (Magnetic resonance imaging, MRI) 中的独特信息[83]. 而更具有吸引力的是, cEViT-GAN 模型还有可能被用作生物标记识别工具, 用于识别人脑的结构和功能连接, 那么这项技术在未来将很有可能进一步用来挖掘 PFC 的结构和内部功能连接情况, 在帮助我们更好地认识 PFC 的同时, 引导我们提出更好的类脑神经网络模型. 另一方面, PFC 也可以启发 Transformer 更好地应用至视觉处理任务中, 例如受认知科学启发, 研究人员提出一种新型双通路 CNN 架构——认知启发网络 (Cognition-inspired network,

CogNet), 用以模仿人类大脑的全局−局部视觉信息处理机制[84]. CogNet 由全局通路、局部通路和自上而下调制器组成, 其中全局通路的形成正是源于在认知科学中发现的一条参与物体识别、更快且途经 PFC 的皮层下通路[85−86], 在此基础上优化标准 Transformer 编码器, 以捕捉输入图像中局部部分的全局结构和上下文信息. 由此可见 PFC 和 Transformer 可以在视觉处理领域中相互促进、共同发展.

3) PFC 也可以为 Transformer 在强化学习 (Reinforcement learning, RL) 领域任务中的发展提供思路. 强化学习是一种机器学习方法, 它通过让智能体在环境中不断尝试不同的动作并根据获得的奖励来调整策略, 以达到最大化累计奖励的目标. 目前在生物领域中, 关于 PFC 与强化学习之间联系的研究已经有一定进展. 例如有研究者利用 fMRI 技术, 通过引入一个任务重映射范式, 让被试者解决多个结构或感官属性不同的 RL 问题, 从而探讨腹内侧前额叶皮层 (Ventromedial prefrontal cortex, vmPFC) 如何抽象和概括强化学习任务的结构[87]. 此外, 奖励机制作为强化学习的重点研究方向, 也有研究者围绕这个方向研究内侧前额叶皮层 (mPFC) 不同亚区在以线索引导的风险/奖励决策中的不同作用, 通过对老鼠进行“二十一点任务”测试分析内侧前额叶皮层的腹侧和背侧区域失活对决策行为的影响, 研究结果表明 mPFC 的不同亚区在风险/奖励决策中具有不同且互补的作用[88]. 还有研究者在对人类和啮齿动物的研究过程中发现 vmPFC 在状态推断过程中发挥着作用, vmPFC 参与预测奖励任务并在人类中显示出与基于状态的推断模型的相关性[89]. 以上这些研究发现表明 PFC 可以为强化学习提供生物层面的解释, 而与此同时 Transformer 模型也被发现可以与强化学习产生密切的联系[90], 因此与强化学习相关的 PFC 和 Transformer 也可以在未来的研究中搭建起桥梁, 通过 PFC 对 Transformer 的启发进一步促进其在强化学习领域的发展.

PFC 启发 Transformer 模型在自然语言处理、视觉处理和强化学习领域中的应用已经展示出强大的潜力和发展前景. 未来, PFC 启发 Transformer 还有望在医疗诊断、金融风险评估等更多领域得到应用和发展, 随着研究的不断深入, 期待它们最终可以为社会发展、科技进步贡献更多力量.

3.2 PFC 启发多类型神经网络模型

本文围绕 PFC 对 Transformer 模型的结合展开一系列的讨论, 那么这也引申出另一个问题: PFC 是否同样可以启发其他诸多神经网络模型呢? 对此进行了进一步的研究.

首先, 将目光聚焦在传统的 CNN 上, PFC 可以与 CNN 在模型架构、算法、应用任务等层面进行结合, 从而启发其进一步地发展. 具体来说, 近年来有学者对识别极端环境下操作员容易出现的异常心理、生理状态进行相关研究, 通过将 CNN 和图注意力网络两种神经网络进行融合, 并从 fNIRS 前额叶皮层网络中提取具有较深神经科学领域知识的鉴别特征, 从而设计出操作员绩效评估模型, 并有效地抑制模型的过拟合现象[91]. 还有研究者对比研究 PFC 活动和 CNN 预测情况, 发现 PFC 信号与 CNN 衍生的决策值 (Decision value, DV) 存在相关性, 且二者在环境选择过程中表现出相似的计算表征. 通过表征相似性分析 (Representational similarity analyses, RSA), PFC 和 CNN 在提取环境信息时有相似的计算机制, 并且研究发现 PFC 与 CNN 的架构特点存在关联, 这些重要的发现为进一步探究大脑在复杂环境选择中的决策机制提供了基础[92]. 此外, 还有研究人员研究 CNN 与人类视觉感知的神经表征, 其中涉及到 PFC 在人类视觉感知中的作用以及 PFC 与 CNN 的关联和对比, 通过比较人类视觉感知和 CNN 模型的神经表征, 实验中发现人类视觉感知涉及到 PFC 等多个脑区, 并且比 CNN 模型的神经表征更广泛地分布在整个大脑中. 这表明 CNN 模型可以从包括 PFC 功能的人类视觉系统中获取灵感, 进一步改进其架构和算法, 以更好地模拟人类的视觉感知和认知过程[93]. 这些 CNN 与 PFC 之间的启发应用与相互结合证明了二者存在的诸多相似之处以及光明的结合发展前景.

其次, 对脉冲神经网络 (Spiking neural network, SNN) 的相关内容进行调研. 在生物领域中, 大脑中的神经元利用动作电位通过突触进行学习和信息传递, SNN 的工作原理与之相似, 即如果输入电压超过阈值便可以产生人工动作电位. SNN 正是由于其与大脑神经元的生物相似性被选择用于计算模型并使用生物学上的方案构建网络结构, 其中包括侧向抑制和使用脉冲时序依赖性可塑性 (Spike timing dependent plasticity, STDP) 规则的自适应突触权重更新[94]. 目前, SNN 已经可以成功地应用于分类和模式识别等问题, 例如利用 SNN 和 STDP 建立模拟哺乳动物嗅觉系统的神经元模型等[95]. 在相关研究中已有的结果表明, 只要涉及到任务转换的过程, 大脑中 PFC 区域的活动就会增加[96−98], 为将所观察到的这种现象通过模型模拟出来, 研究人员使用 SNN 对 PFC 神经元构建一个使用 STDP

作为学习规则的计算模型, 从而可以达到模拟任务切换效果的目的[94]. 还有研究者针对 SNN 和 PFC 中的运动前皮层 (Premotor cortex, PMC) 的结合展开研究[99]. 在该研究中 SNN 被用来模拟 PMC 中的微观神经活动, 并利用加权邻接矩阵与中尺度发射率的乘积来逼近宏观轨迹, 最后通过递归最小二乘方法更新加权邻接矩阵. 以上提到的和类似的应用都体现在面对特定情境时, 可以通过 PFC 中表现出的神经元活动来启发 SNN 更好地对表现出的神经活动进行捕捉和建模.

除此之外, 我们还对 PFC 在图神经网络 (Graph neural network, GNN) 中的应用进行研究. 传统的机器学习和深度学习算法未能通过定位感兴趣区域 (Regions of interest, ROIs) 来提供大脑连接的映射[100], 原因在于这些模型通常忽视大脑的基本网络结构. 图神经网络的优势则在于能够捕捉图域中特征相互作用所包含的结构信息, 可以较好地弥补传统深度学习无法定位出感兴趣区域的不足, 因此近年来在神经科学领域获得极大的关注. 相关研究中, 有研究者基于图卷积网络的分类器, 通过 EEG 记录的工作记忆试验情况, 对健康对照组和带有 PFC 损伤的患者进行群体级分类[101]. 还有生物医学研究者在研究中发现现有的 GNN 模型可能会忽略大脑的局部特异性, 进而在如何使用图结构精确表示复杂的大脑连接组问题上产生困难与挑战. 考虑到功能性紊乱包裹在大脑连接组中的空间分布并不均匀, 因此传统的图学习方法无法完全模拟神经心理疾病的大脑连接组, 面对此问题, 研究者们在最先进的连接组关联研究方法 (Connectome-wide association studies, CWAS) 中多元距离矩阵回归 (Multivariate distance matrix regression, MDMR) 的启发下提出一种基于多变量距离的连接组网络 (Multivariate distance-based connectome network, MDCN)[102], 进而解决无法完全模拟大脑连接组的问题. 上述例子以及类似研究中, 展现了 PFC 更多的是在区域定位信息整合以及复杂网络处理等角度来给予 GNN 启发, 从而结合得到更加具有认知能力的图神经网络, 使其能够更好地处理复杂的空间分布和区域定位等任务.

最后, 我们对 PFC 在循环神经网络 (RNN) 中的应用进行研究. RNN 是一种具有循环连接的神经网络, 它可以处理类似时间序列、文本的序列数据. 目前, RNN 在人工智能领域得到了广泛的应用并且在不断发展和改进, 已经诞生了包含空间嵌入循环神经网络 (Spatially embedded recurrent neural network, seRNN) 在内的诸多 RNN 模型. seRNN 在三维欧几里得空间中存在并收敛于结构和功能特征, 它揭示了许多常见的结构和功能脑模式是如何紧密相互交织并归因于基本的生物优化过程的[103]. 还有研究者通过 PFC 的功能与 RNN 结合, 建立能够在面对新任务或环境时快速适应并通过回放机制不断优化决策过程的模型. 模型中的回放机制类似于海马回放对 PFC 的反馈, 这有助于解释人类行为中的认知需求和行动选择的整体过程[104]. 此外, LSTM 作为 RNN 的一种变体, 引入细胞状态、输入门、遗忘门和输出门等结构, 从而被设计用来解决 RNN 中的梯度消失和梯度爆炸问题, 并使得 LSTM 可以长时间地保存和处理信息, 更加有效地处理长序列数据. 近年来的研究中也诞生了许多 PFC 与 LSTM 相结合的案例. 例如, 有研究者在利用脑电图信号对抑郁症进行检测研究的过程中, 通过对 PFC 相关脑电波数据进行分析处理以及 LSTM 模型的搭建应用, 成功地得到一种有效的抑郁症检测方法[105]. 同样是在抑郁症检测方向, 还有研究者利用双向长短期记忆网络 (Bidirectional long short-term memory, Bi-LSTM) 对脑电图数据进行分类, 并采用最小冗余最大相关性 (Minimum redundancy maximum relevance, mRMR) 方法进行特征选择, 从而研究出基于 PFC 区域 (Fp1、Fpz、Fp2) 脑电图的抑郁症检测系统[106]. 因此, PFC 对 RNN、LSTM 模型的启发主要表现在处理决策制定、执行控制、时间序列数据处理、灵活的动态调节等应用任务方面.

另外, 无论是 CNN、SNN、RNN 中的哪一种人工神经网络模型, 激活函数都在其中起着至关重要的作用, 因此我们同样对 PFC 与激活函数之间的联系进行了部分调研.

近年来, 人工神经网络在学术界和工业界取得显著成功. 激活函数在神经网络的学习过程中起着至关重要的作用. 早期的激活函数包括线性函数、阶跃函数、Sigmoid 函数等[107], 这些函数在神经网络发展的早期阶段被广泛使用. 近年来, 自适应激活函数成为研究热点[108]. 这些函数可以根据网络的训练情况自动调整其参数, 从而提高网络的性能[109]. 未来的研究中, 开发更高效的自适应激活函数、研究激活函数在不同领域的应用、探索激活函数与其他技术的结合等都是值得探索的方向[110]. 有学者通过构建基于组织电导率和人体解剖数据的三维导体模型以及有髓神经纤维模型, 从而发现激活函数可由刺激诱导的电位场对神经纤维膜电压的二阶差分近似表示, 这样的发现对研究包含 PFC 在内的大脑区域在运动皮层刺激中的作用有重要意义[111].

关于激活函数的非线性作用, 有学者从理论和数值结果两方面阐述其对基频分量增强的作用, 同时从经典傅里叶分析角度探讨一般激活函数的选择原则, 体现激活函数的非线性作用在信号处理中的应用及理论依据[112]. 还有学者提出在深度学习应用中, 神经形态光子学旨在满足复杂计算需求, 但现有方法大多关注神经网络的线性部分, 非线性部分的激活函数通常仍在数字电子领域实现. 目前已经有一些模拟电光和全光非线性激活函数被提出, 但仍缺少如 tanh 这样广泛应用的非线性函数的光子学实现及实验验证[113]. 这些相关研究都揭示了非线性在激活函数中的关键作用以及继续深入 PFC 与激活函数之间联系的重要意义.

通过以上内容, 对于本节伊始提出的问题给出了很好的解答, PFC 确实可以为多种类型的神经网络模型提供灵感, 使模型能够更好地模拟 PFC 的结构与功能, 进而在各自的应用领域中实现更加智能和灵活的信息处理.

3.3 PFC 启发 Transformer 构建世界模型

世界模型是一个比较复杂且新颖的概念, 它是对真实世界的一种抽象的表示, 通常用于模拟和理解真实世界的行为与现象. 目前在研究中发现 Transformer、PFC 都可以与世界模型在各自的领域中建立联系, 那么 PFC 是否也有可能启发 Transformer 更好地建立世界模型呢?本节将基于这一值得探究的问题进行展开.

人类智能的显著特征包括高级认知能力和在与世界及自身的各种互动中体现的控制能力, 这些能力并非事先定义的并且会随时间变化. 因此, 类人智能机器的建立以及大脑科学、行为分析、机器人技术及其相关的理论形式化的进展, 凸显世界模型学习和推理的重要性. 在过去几十年, 新的模型学习算法不断取得进展, 这一趋势不仅刺激机器学习研究人员通过构建世界模型来实现类人识别和动作控制[114−115], 而且还启发神经科学研究人员利用这些技术对现实世界中的自然数据进行解码并创建类似于大脑中的处理过程[116]. 可以说, 世界模型是推动自然智能和人工智能协同研究的关键结构, 二者相互结合、和谐发展. 在机器人研究领域, 从概率生成模型的角度来看, 学习世界模型相当于通过推断局部和全局潜变量来模拟可观物理世界[117]. 与此同时, 在机器人领域中实现世界模型学习和推理, 并使其能够在日常环境中进行终身学习, 是一项严峻的挑战, 主要包含机器人本体、多模态信号整合与理解、世界模型建立、自我具身模型建立等多方面的挑战, 这些困难与挑战还需要在进一步的研究中不断解决.

世界模型在强化学习领域有着十分广泛的应用. 最近, 基于模型的强化学习算法在视觉输入环境中表现出显著的效果. 在这些方法中通过自监督学习构建一个参数化的仿真世界模型来模拟真实环境, 并且利用世界模型的想象力增强智能体的选择策略空间, 从而使其不受来自真实环境的采样限制. 这些算法的性能在很大程度上取决于世界模型的序列建模和生成能力, 虽然研究人员已经多次尝试将 Transformer 纳入世界模型之中[118−120], 但在这些工作中并没有充分利用好 Transformer 架构的功能. 对此, 有研究者针对这个问题将 Transformer 的强大序列建模和生成能力与变分自动编码器 (VAE) 的随机特性相结合, 提出一种高效的世界模型架构——基于 Stochastic Transformer 的世界模型 (Stochastic transformer-based world model, STORM)[121]. 它遵循基于模型强化学习算法的既定框架, 选择将重点放在通过想象力增强智能体的策略[119, 122−125], 该框架的基本思路在于通过执行当前策略并将真实环境数据添加到回放缓冲区中, 进而使用从回放缓冲区中采样得到的轨迹更新世界模型, 最后使用得到的世界模型对已有策略进行改进, 通过这样的步骤不断迭代, 直到模型达到指定数量的真实环境交互能力. 因此, 通过与环境交互来构建世界模型的基于模型的强化学习 (Model-based reinforcement learning, MBRL)[126] 已经成为一种很有前途的方法. 作为 MBRL 的基本组成部分, 这些世界模型使人工智能体能够预测其行为的后果并相应地制定计划. 由于时间序列建模的深度神经网络模型快速发展, 为视觉 MBRL 选择合适的主干体系结构已经成为一个相当大的挑战, 特别是 Transformer 和结构化状态空间序列模型 S4[127]、S5[128] 之间的选择. 目前, 已经有很多研究探索将 Transformer 作为世界模型的主干[118−120], 同时也有研究者提出了第一个与可并行 SSM 兼容的通用世界模型框架——S4WM[129], 并通过与视觉世界模型的三种主干架构进行实证比较研究, 得出了 S4WM 在多个记忆要求任务上优于 RNN 和 Transformer 的结论. 这些都体现了世界模型在强化学习领域的重要性以及它所拥有的良好的发展前景, 相信未来通过 PFC、生物领域相关技术与 Transformer 结合以及 Transformer 与强化学习加强联系后, 可以构建出性能更好的世界模型.

世界模型还有望在自动驾驶领域得到突破, 为人类的交通出行带来便利. 最近先进的端到端自动

驾驶方法建立了潜在世界模型, 通过将高维观测抽象为紧凑的潜在状态, 使自动驾驶汽车能够向前预测并从低维状态中学习. 有学者在此背景下提出一种端到端自动驾驶框架, 该框架采用随机顺序潜在世界模型来降低强化学习的样本复杂度问题[130]. 还有研究进一步将具有确定性路径和随机路径的新循环世界模型 dream[123, 131] 引入自动驾驶框架中, 从而提高预测精度和驾驶策略学习过程的稳定性[132]. 然而, 以往工作中提出的世界模型可能仍然包含大量与任务无关的冗余信息, 这会导致低采样效率和对输入扰动的鲁棒性差. 为解决上述问题, 研究者进而提出一种语义遮罩递归世界模型 (Semantic masked recurrent world model, SEM2), 选择通过引入语义过滤器来提取与驾驶相关的关键特征, 并通过过滤后的特征来做出合理的决策[133]. 自动驾驶对于提升交通安全、节约能源资源等方面都有着十分重要的意义, 目前世界模型在这个领域中已经开始初露锋芒, 相信在未来的研究中随着世界模型的不断发展以及诸如 PFC 的生物信息带来启发, 可以推进包含自动驾驶在内的科技领域迸发出更多的创新.

也有研究为生物领域中存在世界模型提供了证明, 生物领域中更多的是体现在神经元与神经元相互连接所形成环境的内部世界模型. 例如有生物研究人员以脉冲神经网络为计算框架, 构建海马和 PFC 神经元及其神经连接回路的放电模型, 进而在大鼠随机探索过程中, 观察到海马结构编码自我运动信息、位置神经元特异性放电等现象[134], 从而证实了这一观点. 既然在生物领域中 PFC 可以与世界模型进行关联, 同时在人工智能领域的 Transformer 中也展现出了与世界模型很好的结合能力, 那么通过 PFC 的相关功能启发 Transformer 更好地与世界模型进行结合就成为一个值得展望的研究方向, 相信在未来的研究中会有更多相关成果的诞生.

3.4 研究展望

随着对 PFC 启发 Transformer 及其他神经网络模型的研究不断深入, 未来充满着无限可能. 一方面, 我们应该进一步加强神经科学、人工智能与其他学科领域的深度融合与知识交流, 拓展 PFC 启发的神经网络模型在更多领域的应用. 另一方面, 还应该继续探索 PFC 与 Transformer 等神经网络模型的结合方式, 优化现有的神经网络模型并开发新型的神经网络模型, 推动世界模型的发展和应用. 同时, 还需要解决数据质量和数量、模型复杂性和计算成本、模型可解释性等方面的技术挑战, 以实现人工智能的更大突破和创新, 为社会发展和科技进步做出更大贡献.

4 结束语

本文通过总结生物学中前额叶皮层的注意力机制、生物编码、多感觉融合等功能, 并与人工智能领域 Transformer 模型的自注意力机制、位置编码、多模态能力进行一一对应, 从而挖掘出前额叶皮层启发 Transformer 模型已有的研究进展以及未来可行的研究思路. 伴随着前额叶皮层和 Transformer 在各自领域的不断发展, 相信未来会出现更多生物结构与功能启发人工智能模型的新发现, 也相信我们终将会搭建起连接生物与人工智能两大领域的类脑计算桥梁.

References

1 Ceccarelli F, Ferrucci L, Londei F, Ramawat S, Brunamonti E, Genovesio A. Static and dynamic coding in distinct cell types during associative learning in the prefrontal cortex. *Nature Communications*, 2023, **14**(1): Article No. 8325

2 Watakabe A, Skibbe H, Nakae K, Abe H, Ichinohe N, Rachmadi M F, et al. Local and long-distance organization of prefrontal cortex circuits in the marmoset brain. *Neuron*, 2023, **111**(14): 2258−2273

3 Passingham R E, Lau H. Do we understand the prefrontal cortex? *Brain Structure and Function*, 2023, **228**(5): 1095−1105

4 Trapp N T, Bruss J E, Manzel K, Grafman J, Tranel D, Boes A D. Large-scale lesion symptom mapping of depression identifies brain regions for risk and resilience. *Brain*, 2023, **146**(4): 1672−1685

5 Chafee M V, Heilbronner S R. Prefrontal cortex. *Current Biology*, 2022, **32**(8): R346−R351

6 Miller E K, Cohen J D. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 2001, **24**: 167−202

7 Diehl G W, Redish A D. Differential processing of decision information in subregions of rodent medial prefrontal cortex. *eLife*, 2023, **12**: Article No. e82833

8 Wang J X, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo J Z, et al. Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, 2018, **21**(6): 860−868

9 Alexander W H, Brown J W. Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, 2011, **14**(10): 1338−1344

10 Rikhye R V, Gilra A, Halassa M M. Thalamic regulation of switching between cortical representations enables cognitive flexibility. *Nature Neuroscience*, 2018, **21**(12): 1753−1763

11 Gisiger T, Boukadoum M. Mechanisms gating the flow of information in the cortex: What they might look like and what their uses may be. *Frontiers in Computational Neuroscience*, 2011, **5**: Article No. 1

12 Johnston K, Levin H M, Koval M J, Everling S. Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. *Neuron*, 2007, **53**(3): 453−462

13 Tsuda B, Tye K M, Siegelmann H T, Sejnowski T J. A modeling framework for adaptive lifelong learning with transfer and savings through gating in the prefrontal cortex. *Proceedings of*

the National Academy of Sciences of the United States of America, 2020, **117**(47): 29872−29882

14 Wang Z H, Zhang J, Zhang X C, Chen P, Wang B. Transformer model for functional near-infrared spectroscopy classification. *IEEE Journal of Biomedical and Health Informatics*, 2022, **26**(6): 2559−2569

15 Choi S R, Lee M. Transformer architecture and attention mechanisms in genome data analysis: A comprehensive review. *Biology*, 2023, **12**(7): Article No. 1033

16 Li Q P, Zhuang Y. An efficient image-guided-based 3D point cloud moving object segmentation with transformer-attention in autonomous driving. *International Journal of Applied Earth Observation and Geoinformation*, 2023, **123**: Article No. 103488

17 Brus J, Heng J A, Beliaeva V, Gonzalez Pinto F, Cassarà A M, Neufeld E, et al. Causal phase-dependent control of non-spatial attention in human prefrontal cortex. *Nature Human Behaviour*, 2024, **8**(4): 743−757

18 Bichot N P, Heard M T, Degennaro E M, Desimone R. A source for feature-based attention in the prefrontal cortex. *Neuron*, 2015, **88**(4): 832−844

19 Huang L, Wang J Y, He Q H, Li C, Sun Y L, Seger C A, et al. A source for category-induced global effects of feature-based attention in human prefrontal cortex. *Cell Reports*, 2023, **42**(9): Article No. 113080

20 Zhao M L, Xu D H, Gao T. From cognition to computation: A comparative review of human attention and transformer architectures. arXiv preprint arXiv: 2407.01548, 2024.

21 Kumar S, Sumers T R, Yamakoshi T, Goldstein A, Hasson U, Norman K A, et al. Shared functional specialization in transformer-based language models and the human brain. *Nature Communications*, 2024, **15**(1): Article No. 5523

22 Muller L, Churchland P S, Sejnowski T J. Transformers and cortical waves: Encoders for pulling in context across time. *Trends in Neurosciences*, 2024, **47**(10): 788−802

23 Huang H M, Li R, Qiao X J, Li X R, Li Z Y, Chen S Y, et al. Attentional control influence habituation through modulation of connectivity patterns within the prefrontal cortex: Insights from stereo-EEG. *NeuroImage*, 2024, **294**: Article No. 120640

24 Li N N, Chen Y R, Li W F, Ding Z X, Zhao D B, Nie S. BViT: Broad attention-based vision transformer. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, **35**(9): 12772−12783

25 Shi Q Q, Fan J S, Wang Z R, Zhang Z X. Multimodal channel-wise attention transformer inspired by multisensory integration mechanisms of the brain. *Pattern Recognition*, 2022, **130**: Article No. 108837

26 Gong D Y, Zhang H T. Self-attention limits working memory capacity of transformer-based models. arXiv preprint arXiv: 2409.10715, 2024.

27 Maith O, Schwarz A, Hamker F H. Optimal attention tuning in a neuro-computational model of the visual cortex-basal ganglia-prefrontal cortex loop. *Neural Networks*, 2021, **142**: 534−547

28 Spitale G, Biller-Andorno N, Germani F. AI model GPT-3 (dis) informs us better than humans. *Science Advances*, 2023, **9**(26): Article No. eadh1850

29 Brown T B, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2020. Article No. 159

30 Zhang S S, Roller S, Goyal N, Artetxe M, Chen M Y, Chen S H, et al. OPT: Open pre-trained transformer language models. arXiv preprint arXiv: 2205.01068, 2022.

31 Yue F P, Ko T. An investigation of positional encoding in transformer-based end-to-end speech recognition. In: Proceedings of the 12th International Symposium on Chinese Spoken Language Processing (ISCSLP). Hong Kong, China: IEEE, 2021. 1−5

32 Kazemnejad A, Padhi I, Natesan Ramamurthy K, Das P, Reddy S. The impact of positional encoding on length generalization in transformers. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc., 2024. Article No. 1082

33 Chowdhery A, Narang S, Devlin J, Bosma M, Mishra G, Roberts A, et al. PaLM: Scaling language modeling with pathways. *Journal of Machine Learning Research*, 2023, **24**(240): 1−113

34 Zhang R R, Han J M, Liu C, Gao P, Zhou A J, Hu X F, et al. LLaMA-adapter: Efficient fine-tuning of language models with zero-init attention. arXiv preprint arXiv: 2303.16199, 2023.

35 Touvron H, Lavril T, Izacard G, Martinet X, Lachaux M A, Lacroix T, et al. LLaMA: Open and efficient foundation language models. arXiv preprint arXiv: 2302.13971, 2023.

36 Wu J S, Zhang R C, Mao Y Y, Chen J F. On scalar embedding of relative positions in attention models. In: Proceedings of the 35th AAAI Conference on Artificial Intelligence. AAAI, 2021. 14050−14057

37 Wallis J D, Anderson K C, Miller E K. Single neurons in prefrontal cortex encode abstract rules. *Nature*, 2001, **411**(6840): 953−956

38 Oota S R, Arora J, Rowtula V, Gupta M, Bapi R S. Visio-linguistic brain encoding. In: Proceedings of the 29th International Conference on Computational Linguistics. Gyeongju, Republic of Korea: ACL, 2022. 116−133

39 Bocincova A, Buschman T J, Stokes M G, Manohar S G. Neural signature of flexible coding in prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 2022, **119**(40): Article No. e2200400119

40 Zhang K, Hao W N, Yu X H, Shao T H. An interpretable image classification model combining a fuzzy neural network with a variational autoencoder inspired by the human brain. *Information Sciences*, 2024, **661**: Article No. 119885

41 Aoi M C, Mante V, Pillow J W. Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nature Neuroscience*, 2020, **23**(11): 1410−1420

42 Zhang Z M, Gong X. Positional label for self-supervised vision transformer. In: Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington, USA: AAAI, 2023. 3516−3524

43 Caucheteux C, Gramfort A, King J R. Evidence of a predictive coding hierarchy in the human brain listening to speech. *Nature Human Behaviour*, 2023, **7**(3): 430−441

44 Busch A, Roussy M, Luna R, Leavitt M L, Mofrad M H, Gulli R A, et al. Neuronal activation sequences in lateral prefrontal cortex encode visuospatial working memory during virtual navigation. *Nature Communications*, 2024, **15**(1): Article No. 4471

45 Labaien J, Idé T, Chen P Y, Zugasti E, de Carlos X. Diagnostic spatio-temporal transformer with faithful encoding. *Knowledge-Based Systems*, 2023, **274**: Article No. 110639

46 Deihim A, Alonso E, Apostolopoulou D. STTRE: A spatio-temporal transformer with relative embeddings for multivariate time series forecasting. *Neural Networks*, 2023, **168**: 549−559

47 Ma Y J, Wang R L. Relative-position embedding based spatially and temporally decoupled Transformer for action recognition. *Pattern Recognition*, 2024, **145**: Article No. 109905

48 Coen P, Sit T P H, Wells M J, Carandini M, Harris K D. Mouse frontal cortex mediates additive multisensory decisions. *Neuron*, 2023, **111**(15): 2432−2447

49 Ferrari A, Noppeney U. Attention controls multisensory perception via two distinct mechanisms at different levels of the cortical hierarchy. *PLoS Biology*, 2021, **19**(11): Article No.

e3001465

50 Mihalik A, Noppeney U. Causal inference in audiovisual perception. *The Journal of Neuroscience*, 2020, **40**(34): 6600−6612

51 Kang K, Rosenkranz R, Karan K, Altinsoy E, Li S C. Congruence-based contextual plausibility modulates cortical activity during vibrotactile perception in virtual multisensory environments. *Communications Biology*, 2022, **5**(1): Article No. 1360

52 Cao Y N, Summerfield C, Park H, Giordano B L, Kayser C. Causal inference in the multisensory brain. *Neuron*, 2019, **102**(5): 1076−1087

53 Giessing C, Thiel C M, Stephan K E, Rösler F, Fink G R. Visuospatial attention: How to measure effects of infrequent, unattended events in a blocked stimulus design. *Neuroimage*, 2004, **23**(4): 1370−1381

54 Zheng Q H, Zhou L X, Gu Y. Temporal synchrony effects of optic flow and vestibular inputs on multisensory heading perception. *Cell Reports*, 2021, **37**(7): Article No. 109999

55 Liang P P, Zadeh A, Morency L P. Foundations & trends in multimodal machine learning: Principles, challenges, and open questions. *ACM Computing Surveys*, 2024, **56**(10): Article No. 264

56 Klemen J, Chambers C D. Current perspectives and methods in studying neural mechanisms of multisensory interactions. *Neuroscience and Biobehavioral Reviews*, 2012, **36**(1): 111−133

57 Paraskevopoulos G, Georgiou E, Potamianos A. Mmlatch: Bottom-up top-down fusion for multimodal sentiment analysis. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Singapore: IEEE, 2022. 4573−4577

58 Sun H, Liu J Q, Chen Y W, Lin L F. Modality-invariant temporal representation learning for multimodal sentiment classification. *Information Fusion*, 2023, **91**: 504−514

59 Wang Z L, Wan Z H, Wan X J. TransModality: An End2End fusion method with transformer for multimodal sentiment analysis. In: Proceedings of the Web Conference. Taipei, China: ACM, 2020. 2514−2520

60 Yu J F, Chen K, Xia R. Hierarchical interactive multimodal transformer for aspect-based multimodal sentiment analysis. *IEEE Transactions on Affective Computing*, 2023, **14**(3): 1966−1978

61 Huang J H, Zhou J, Tang Z C, Lin J Y, Chen C Y C. TMBL: Transformer-based multimodal binding learning model for multimodal sentiment analysis. *Knowledge-Based Systems*, 2024, **285**: Article No. 111346

62 Yang D K, Liu Y, Huang C, Li M C, Zhao X, Wang Y Z, et al. Target and source modality co-reinforcement for emotion understanding from asynchronous multimodal sequences. *Knowledge-Based Systems*, 2023, **265**: Article No. 110370

63 Ahn H J, Lee D H, Jeong J H, Lee S W. Multiscale convolutional transformer for EEG classification of mental imagery in different modalities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023, **31**: 646−656

64 Li J, Chen N, Zhu H Q, Li G Q, Xu Z Y, Chen D X. Incongruity-aware multimodal physiology signals fusion for emotion recognition. *Information Fusion*, 2024, **105**: Article No. 102220

65 Asif M, Gupta A, Aditya A, Mishra S, Tiwary U S. Brain multi-region information fusion using attentional transformer for EEG based affective computing. In: Proceedings of the IEEE 20th India Council International Conference (INDICON). Hyderabad, India: IEEE, 2023. 771−775

66 Chen Z H, Han Y C, Ma Z, Wang X N, Xu S R, Tang Y, et al. A prefrontal-thalamic circuit encodes social information for social recognition. *Nature Communications*, 2024, **15**(1): Article No. 1036

67 Yu J, Li J, Yu Z, Huang Q M. Multimodal transformer with multi-view visual representation for image captioning. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, **30**(12): 4467−4480

68 Hu B, Guan Z H, Chen G R, Chen C L P. Neuroscience and network dynamics toward brain-inspired intelligence. *IEEE Transactions on Cybernetics*, 2022, **52**(10): 10214−10227

69 Sucholutsky I, Muttenthaler L, Weller A, Peng A D, Bobu A, Kim B, et al. Getting aligned on representational alignment. arXiv preprint arXiv: 2310.13018, 2023.

70 Chersoni E, Santus E, Huang C R, Lenci A. Decoding word embeddings with brain-based semantic features. *Computational Linguistics*, 2021, **47**(3): 663−698

71 Toneva M, Wehbe L. Interpreting and improving natural-language processing (in machines) with natural language-processing (in the brain). In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2019. Article No. 1339

72 Yu S Y, Gu C Y, Huang K X, Li P. Predicting the next sentence (not word) in large language models: What model-brain alignment tells us about discourse comprehension. *Science Advances*, 2024, **10**(21): Article No. eadn7744

73 Cambria E, Das D, Bandyopadhyay S, Feraco A. Affective computing and sentiment analysis. *A Practical Guide to Sentiment Analysis*. Cham: Springer, 2017. 1−10

74 Mishra A, Dey K, Bhattacharyya P. Learning cognitive features from gaze data for sentiment and sarcasm classification using convolutional neural network. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver, Canada: ACL, 2017. 377−387

75 Zhang Z H, Wu C H, Chen H Y, Chen H Y. CogAware: Cognition-aware framework for sentiment analysis with textual representations. *Knowledge-Based Systems*, 2024, **299**: Article No. 112094

76 Montejo-Ráez A, Molina-González M D, Jiménez-Zafra S M, García-Cumbreras M Á, García-López L J. A survey on detecting mental disorders with natural language processing: Literature review, trends and challenges. *Computer Science Review*, 2024, **53**: Article No. 100654

77 Ramachandran G, Yang R. CortexCompile: Harnessing cortical-inspired architectures for enhanced multi-agent NLP code synthesis. arXiv preprint arXiv: 2409.02938, 2024.

78 Li Z J, Zhao B, Zhang G Y, Dang J W. Brain network features differentiate intentions from different emotional expressions of the same text. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island, Greece: IEEE, 2023. 1−5

79 Squires M, Tao X H, Elangovan S, Gururajan R, Zhou X J, Acharya U R, et al. Deep learning and machine learning in psychiatry: A survey of current progress in depression detection, diagnosis and treatment. *Brain Informatics*, 2023, **10**(1): Article No. 10

80 Song G Z, Huang D G, Xiao Z F. A study of multilingual toxic text detection approaches under imbalanced sample distribution. *Information*, 2021, **12**(5): Article No. 205

81 Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, et al. An image is worth 16 × 16 words: Transformers for image recognition at scale. arXiv preprint arXiv: 2021.11929, 2020.

82 Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc., 2017. 6000−6010

83 Bi Y D, Abrol A, Jia S H, Fu Z N, Calhoun V D. Gray matters: An efficient vision transformer GAN framework for predicting functional network connectivity biomarkers from brain structure. BioRxiv, 2024.

84 Dong S L, Gong Y H, Shi J G, Shang M, Tao X Y, Wei X, et al. Brain cognition-inspired dual-pathway CNN architecture for image classification. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, **35**(7): 9900−9914

85 Liu L, Wang F, Zhou K, Ding N, Luo H. Perceptual integration rapidly activates dorsal visual pathway to guide local processing in early visual areas. *PLoS Biology*, 2017, **15**(11): Article No. e2003646

86 Bar M. The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 2007, **11**(7): 280−289

87 Baram A B, Muller T H, Nili H, Garvert M M, Behrens T E J. Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. *Neuron*, 2021, **109**(4): 713−723

88 van Holstein M, Floresco S B. Dissociable roles for the ventral and dorsal medial prefrontal cortex in cue-guided risk/reward decision making. *Neuropsychopharmacology*, 2020, **45**(4): 683−693

89 Averbeck B, O′Doherty J P. Reinforcement-learning in fronto-striatal circuits. *Neuropsychopharmacology*, 2022, **47**(1): 147−162

90 Hu S C, Shen L, Zhang Y, Chen Y X, Tao D C. On transforming reinforcement learning with transformers: The development trajectory. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, **46**(12): 8580−8599

91 Zhang Y, Jia M, Chen T, Li M, Wang J Y, Hu X M, et al. A neuroergonomics model for evaluating nuclear power plants operators′ performance under heat stress driven by ECG time-frequency spectrums and fNIRS prefrontal cortex network: A CNN-GAT fusion model. *Advanced Engineering Informatics*, 2024, **62**: Article No. 102563

92 Law C K, Kolling N, Chan C C H, Chau B K H. Frontopolar cortex represents complex features and decision value during choice between environments. *Cell Reports*, 2023, **42**(6): Article No. 112555

93 Lee J, Jung M, Lustig N, Lee J H. Neural representations of the perception of handwritten digits and visual objects from a convolutional neural network compared to humans. *Human Brain Mapping*, 2023, **44**(5): 2018−2038

94 Viswanathan K A, Mylavarapu G, Chen K, Thomas J P. A study of prefrontal cortex task switching using spiking neural networks. In: Proceedings of the 12th International Conference on Advanced Computational Intelligence (ICACI). Dali, China: IEEE, 2020. 199−206

95 Li B Z, Pun S H, Feng W, Vai M I, Klug A, Lei T C. A spiking neural network model mimicking the olfactory cortex for handwritten digit recognition. In: Proceedings of the 9th International IEEE/EMBS Conference on Neural Engineering (NER). San Francisco, USA: IEEE, 2019. 1167−1170

96 Hyafil A, Summerfield C, Koechlin E. Two mechanisms for task switching in the prefrontal cortex. *The Journal of Neuroscience*, 2009, **29**(16): 5135−5142

97 Kushleyeva Y, Salvucci D D, Lee F J. Deciding when to switch tasks in time-critical multitasking. *Cognitive Systems Research*, 2005, **6**(1): 41−49

98 Brass M, von Cramon D Y. The role of the frontal cortex in task preparation. *Cerebral Cortex*, 2002, **12**(9): 908−914

99 Wei Q L, Han L Y, Zhang T L. Learning and controlling multiscale dynamics in spiking neural networks using recursive least square modifications. *IEEE Transactions on Cybernetics*, 2024, **54**(8): 4603−4616

100 Demir A, Koike-Akino T, Wang Y, Haruna M, Erdogmus D. EEG-GNN: Graph neural networks for classification of electroencephalogram (EEG) signals. In: Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Mexico, Mexico: IEEE, 2021. 1061−1067

101 Balaji S S, Parhi K K. Classifying subjects with PFC lesions from healthy controls during working memory encoding via graph convolutional networks. In: Proceedings of the 11th International IEEE/EMBS Conference on Neural Engineering (NER). Baltimore, USA: IEEE, 2023. 1−4

102 Yang Y W, Ye C F, Ma T. A deep connectome learning network using graph convolution for connectome-disease association study. *Neural Networks*, 2023, **164**: 91−104

103 Achterberg J, Akarca D, Strouse D J, Duncan J, Astle D E. Spatially embedded recurrent neural networks reveal widespread links between structural and functional neuroscience findings. *Nature Machine Intelligence*, 2023, **5**(12): 1369−1381

104 Jensen K T, Hennequin G, Mattar M G. A recurrent network model of planning explains hippocampal replay and human behavior. *Nature Neuroscience*, 2024, **27**(7): 1340−1348

105 Pratiwi M. Comparative analysis of brain waves for EEG-based depression detection in the prefrontal cortex lobe using LSTM. In: Proceedings of the 7th International Conference on New Media Studies (CONMEDIA). Bali, Indonesia: IEEE, 2023. 173−178

106 Pratiwi M. EEG-based depression detection in the prefrontal cortex lobe using mRMR feature selection and bidirectional LSTM. *Ultima Computing: Jurnal Sistem Komputer*, 2023, **15**(2): 71−78

107 Sharma S, Sharma S, Athaiya A. Activation functions in neural networks. *International Journal of Engineering Applied Sciences and Technology*, 2020, **4**(12): 310−316

108 Jagtap A D, Kawaguchi K, Karniadakis G E. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 2020, **404**: Article No. 109136

109 Abbasi J, Andersen P Ø. Physical activation functions (PAFs): An approach for more efficient induction of physics into physics-informed neural networks (PINNs). *Neurocomputing*, 2024, **608**: Article No. 128352

110 Jagtap A D, Karniadakis G E. How important are activation functions in regression and classification? A survey, performance comparison, and future directions. *Journal of Machine Learning for Modeling and Computing*, 2023, **4**(1): 21−75

111 Manola L, Roelofsen B H, Holsheimer J, Marani E, Geelen J. Modelling motor cortex stimulation for chronic pain control: Electrical potential field, activating functions and responses of simple nerve fibre models. *Medical and Biological Engineering and Computing*, 2005, **43**(3): 335−343

112 Steinerberger S, Wu H T. Fundamental component enhancement via adaptive nonlinear activation functions. *Applied and Computational Harmonic Analysis*, 2023, **63**: 135−143

113 Pappas C, Kovaios S, Moralis-Pegios M, Tsakyridis A, Giamougiannis G, Kirtas M, et al. Programmable tanh-, ELU-, sigmoid-, and sin-based nonlinear activation functions for neuromorphic photonics. *IEEE Journal of Selected Topics in Quantum Electronics*, 2023, **29**(6: Photonic Signal Processing): Article No. 6101210

114 Ha D, Schmidhuber J. World models. arXiv preprint arXiv: 1803.10122, 2018.

115 Eslami S M A, Jimenez Rezende D, Besse F, Viola F, Morcos A S, Garnelo M, et al. Neural scene representation and rendering. *Science*, 2018, **360**(6394): 1204−1210

116 Yamins D L K, Hong H, Cadieu C F, Solomon E A, Seibert D, Dicarlo J J. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, **111**(23): 8619−8624

117 Friston K, Moran R J, Nagai Y, Taniguchi T, Gomi H, Tenenbaum J. World model learning and inference. *Neural Networks*, 2021, **144**: 573−590

118 Robine J, HöFtmann M, Uelwer T, Harmeling S. Transformer-based world models are happy with 100k interactions. In: Proceedings of the Eleventh International Conference on Learning Representations. Kigali, Rwanda: ICLR, 2023.

119 Micheli V, Alonso E, Fleuret F. Transformers are sample-efficient world models. In: Proceedings of the Eleventh International Conference on Learning Representations. Kigali, Rwanda: ICLR, 2023.

120 Chen C, Wu Y F, Yoon J, Ahn S. TransDreamer: Reinforcement learning with transformer world models. arXiv preprint arXiv: 2202.09481, 2022.

121 Zhang W P, Wang G, Sun J, Yuan Y T, Huang G. STORM: Efficient stochastic transformer based world models for reinforcement learning. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc., 2024. Article No. 1182

122 Hafner D, Pasukonis J, Ba J, Lillicrap T. Mastering diverse domains through world models. arXiv preprint arXiv: 2301.04104, 2023.

123 Hafner D, Lillicrap T P, Norouzi M, Ba J. Mastering Atari with discrete world models. arXiv preprint arXiv: 2010.02193, 2020.

124 Barto A G, Sutton R S, Anderson C W. Looking back on the actor-critic architecture. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, **51**(1): 40−50

125 Kaiser L, Babaeizadeh M, Miłos P, Osiński B, Campbell R H, Czechowski K, et al. Model based reinforcement learning for Atari. In: Proceedings of the 8th International Conference on Learning Representations. Addis Ababa, Ethiopia: ICLR, 2020.

126 Moerland T M, Broekens J, Plaat A, Jonker C M. Model-based reinforcement learning: A survey. *Foundations and Trends® in Machine Learning*, 2023, **16**(1): 1−118

127 Gu A, Goel K, Ré C. Efficiently modeling long sequences with structured state spaces. arXiv preprint arXiv: 2111.00396, 2021.

128 Smith J T H, Warrington A, Linderman S. Simplified state space layers for sequence modeling. In: Proceedings of the Eleventh International Conference on Learning Representations. Kigali, Rwanda: ICLR, 2023.

129 Deng F, Park J, Ahn S. Facing off world model backbones: RNNs, transformers, and S4. In: Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc., 2024. Article No. 3188

130 Chen J Y, Li S E, Tomizuka M. Interpretable end-to-end urban autonomous driving with latent deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, 2022, **23**(6): 5068−5078

131 Hafner D, Lillicrap T, Ba J, Norouzi M. Dream to control: Learning behaviors by latent imagination. In: Proceedings of the 8th International Conference on Learning Representations. Addis Ababa, Ethiopia: ICLR, 2020.

132 Zhang Y H, Mu Y, Yang Y J, Guan Y, Li S E, Sun Q, et al. Steadily learn to drive with virtual memory. arXiv preprint arXiv: 2102.08072, 2021.

133 Gao Z Y, Mu Y, Chen C, Duan J L, Luo P, Lu Y F, et al. Enhance sample efficiency and robustness of end-to-end urban autonomous driving via semantic masked world model. *IEEE Transactions on Intelligent Transportation Systems*, 2024, **25**(10): 13067−13079

134 Yu N G, Lv Z X, Yan J H, Wang Z X. Spatial cognition and decision model based on hippocampus-prefrontal cortex interaction. In: Proceedings of the China Automation Congress (CAC). Chongqing, China: IEEE, 2023. 3754−3759

潘雨辰 北京工业大学信息科学技术学院硕士研究生, 中国科学院脑科学与智能技术卓越创新中心联合培养学生. 2019 年获得北京工业大学工学学士学位. 主要研究方向为类脑模型算法.

E-mail: 18201335023@sina.cn

(**PAN Yu-Chen** Master student at the School of Information Science and Technology, Beijing University of Technology, co-supervised by the Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences. He received his bachelor degree in engineering from Beijing University of Technology in 2019. His main research interest is brain-inspired algorithms.)

贾克斌 博士, 北京工业大学信息科学技术学院教授. 主要研究方向为图像/视频处理技术与生物医学信息处理技术.

E-mail: kebinj@bjut.edu.cn

(**JIA Ke-Bin** Ph.D., professor at the School of Information Science and Technology, Beijing University of Technology. His research interest covers image/video processing technology and biomedical information processing technology.)

张铁林 中国科学院脑科学与智能技术卓越创新中心研究员. 主要研究方向为类脑脉冲神经网络算法, 类脑芯片及 AI for Neuroscience 研究. 本文通信作者.

E-mail: zhangtielin@ion.ac.cn

(**ZHANG Tie-Lin** Professor at the Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences. He research interest covers research of brain-inspired spiking neural network algorithms, brain-inspired chips, and AI for Neuroscience. Corresponding author of this paper.)