

# 评人工智能如何走向新阶段？

陆首群

2022. 5. 5

2019年8月8日，COPU提出“人工智能向何处去？”（探索人工智能的发展方向），同时构建“人工智能国内外跟帖留言平台”，期望从广大科学家、工程师在发展人工智能理论和实践中提炼出来的多元言论，探索人工智能发展方向问题。迄今COPU已收到人工智能国内外跟帖留言926条。我们先后出版了人工智能国内外跟帖留言第一集～第十一集（1～837条跟帖），现出版第十二集（838条～926条），供专家和读者们参考。

全部十二集人工智能国内外跟帖留言包括下列六个方面：①机器学习、深度学习和强化学习（弱人工智能），②可解释性机器学习，③异步脉冲神经网络+神经拟态计算系统，④脑机接口，⑤大规模语义网络（知识图谱）+认知智能，⑥通用人工智能。

在2016条国内外研发AI的跟帖中，记录了AI由弱人工智能向强人工智能转化的过程，在可解释性机器学习的跟帖中，COPU与IBM人工智能研究所展开了深入讨论，在异步脉冲神经网络+神经拟态计算系统的跟帖中，COPU与研发者英特尔的科学家也展开了深入讨论。

在这里，补充提一下脑机接口无创连接技术：

2016年，我国发射天宫二号实验室，天津脑科学中心、天津大学对航天员实行无创脑机接口技术。《nature》杂志主刊为此发表“自然

聚焦——中国脑科学”栏目系列：“脑机接口——梦想之光照进现实生活”。

今年4月30日，东南大学、华南理工大学、北京大学信息科学技术学院、新加坡国立大学、广州琶洲实验室、英国格拉斯哥大学合作组成研发团队，通过非侵入性脑机超平台，成功实现人脑直接无线通信。

## 国内外 AI 跟帖留言 (838-926)

**838**, 通过学习何时行动来实现可解释的强化学习

Alexis Jacq 等, 2022. 3. 16,

谷歌研究, 法国国家信息与自动化研究所, 里尔大学

简介: 传统上, 强化学习 (RL) 旨在决定如何对人工智能体进行优化。我们认为, 决定何时采取行动同样重要。作为人类, 当情况需要时, 我们会从默认的、本能的或记忆中的行为转变为专注的、深思熟虑的行为。为了增强具有这种能力的 RL 代理, 我们建议增强标准马尔可夫决策过程, 并提供一种新的行动模式: 懒惰, 这将决策推迟到默认策略。此外, 我们惩罚非懒惰行为, 以鼓励最小的努力, 并让代理只关注关键决策。我们将由此产生的形式主义命名为懒惰的 MDP。我们研究了惰性 MDP 的理论性质, 表达了值函数并刻画了最优解。然后, 我们根据经验证明, 在惰性 MDP 中学习的策略通常具有某种形式的可解释性: 通过构造, 它们向我们展示了代理控制默认策略的状态。我们认为这些状态和相应的操作很重要, 因为它们解释了默认策略和新的懒惰策略之间的性能差异。以次优策略作为默认策略 (预训练或随机), 我们观察到, 在 Atari 游戏中, 代理能够获得竞争性性能, 同时只在有限的状态子集中进行控制。

**839**, 强化学习中的解释性: 视角与立场

Agneza Krajna 等, 2022. 3. 22,

IEEE

简介: 人工智能 (AI) 已经嵌入到人们日常生活的许多方面, 让人工智能为人们

做决策已经变得很正常。相对于其他机器学习范式，强化学习（RL）模型增加了可解问题的空间。一些最有趣的应用是在预期回报函数不可微的情况下，在未知或未定义的环境中运行，以及在算法发现方面，超过任何教师的表现，代理通过简单的反馈从实验经验中学习。应用范围及其社会影响非常广泛，仅举几个例子：基因组学、博弈（国际象棋、围棋等）、一般优化、金融投资、政府政策、自动驾驶汽车、推荐系统等。因此，通过解释来提高基于 RL 的系统的信任度和透明度至关重要。大多数关于人工智能中可解释性的文章都提供了与监督学习相关的方法，而在 RL 领域中很少有文章涉及这一点。造成这种情况的原因是信用分配问题、奖励延迟，以及无法假设数据是独立且相同分布的（i. i. d.）。本文试图系统地概述可解释 RL 领域的现有方法，并在现有方法的基础上提出一种新的统一分类法。立场部分描述了如何观察可解释性的语用方面。特别强调了接受解释和产生解释的各方之间的差距。为了缩小差距，实现解释的诚实和真实性，我们建立了三个支柱：主动性、风险态度和认识论约束。为此，我们对最短路径问题的简单变体进行了说明。

#### **840, XAI 的信任和依赖——区分态度和行为测量**

Nicolas Scharwski 等，2022. 3. 23，

巴塞尔大学

简介：信任经常被认为是有效使用和实际部署人工智能的一个基本标准。研究人员认为，人工智能应该更加透明，以增加信任，使透明度成为 XAI 的主要目标之一。然而，关于透明度对信任的影响，关于这一主题的实证研究并不确定。对这种模糊性的一种解释可能是，信任在 XAI 中的运作方式不同。在这份立场文件中，

我们主张明确区分依赖的行为（客观）衡量标准和信任的态度（主观）衡量标准。然而，研究人员在试图获取信任时，有时似乎会使用行为测量，尽管态度测量更合适。基于过去的研究，我们强调，将信任和依赖分开是有充分的理论依据的。正确区分这两个概念可以更全面地理解透明度如何影响信任和依赖，从而有利于未来的 XAI 研究。

#### 841, 具有时空压缩和突触卷积阻滞的超低潜伏期脉冲神经网络

Changqing Xu 等, 2022. 3. 18,

西安电子科技大学

简介：尖峰神经网络（Spiking neural networks, SNN）是一种受大脑启发的模型，具有时空信息处理能力强、功耗低、生物合理性高等特点。有效的时空特征使其适用于事件流分类。然而，神经形态数据集，如 N-MNIST、CIFAR10-DVS、DVS128 手势，需要将单个事件聚合为具有更高时间分辨率的帧，用于事件流分类，这会导致较高的训练和推理延迟。在这项工作中，我们提出了一种时空压缩方法，将单个事件聚合为突触电流的几个时间步，以减少训练和推理延迟。为了在高压比下保持 SNN 的准确性，我们还提出了一种突触卷积块来平衡相邻时间步之间的剧烈变化。并引入了具有可学习膜时间常数的多阈值漏积分与点火（LIF），以提高其信息处理能力。我们在神经形态 N-MNIST、CIFAR10-DVS、DVS128 手势数据集上评估了所提出的事件流分类任务方法。实验结果表明，我们提出的方法在几乎所有数据集上都优于最先进的精度，使用的时间步长更少。

## 842, 基于神经元膜电位的无监督概率 Hebbian 可塑性规则

NikhilGarg 等, 2022. 3. 21,

舍布鲁克大学, 里尔大学, 土伦大学, 滑铁卢大学, 巴黎萨克雷大学

简介: 本研究提出了电压依赖性突触可塑性 (VDSP), 这是一种新的大脑启发的无监督局部学习规则, 用于在神经形态硬件上在线实现 Hebb 的可塑性机制。建议的 VDSP 学习规则仅更新突触后神经元棘突上的突触电导, 这将标准的棘突时间依赖性可塑性 (STDP) 更新次数减少两倍。这种更新依赖于突触前神经元的膜电位, 该电位作为神经元实现的一部分随时可用, 因此不需要额外的存储内存。此外, 这种更新还可以调节突触重量, 防止重复刺激时重量爆炸或消失。通过严格的数学分析得出 VDSP 和 STDP 之间的等价性。为了验证 VDSP 的系统级性能, 我们训练了一个用于手写数字识别的单层尖峰神经网络(SNN)。我们报告了 MNIST 数据集上 100 个输出神经网络的准确率为  $85.01 \pm 0.76\%$  (平均值  $\pm$  标准差)。当扩展网络规模时, 性能有所提高 (400 个输出神经元为  $89.93 \pm 0.41\%$ , 500 个神经元为  $90.56 \pm 0.27\%$ ), 这验证了所提出的学习规则对大规模计算机视觉任务的适用性。有趣的是, 学习规则比 STDP 更好地适应输入信号的频率, 并且不需要手动调整超参数。

## 843, 稀疏主动卷积尖峰神经网络的高效硬件加速

Jan Sommer 等, (2022. 3. 23), IEEE

简介: 尖峰神经网络 (SNN) 在基于事件的情况下进行计算, 以实现比标准神经网络更高效的计算。在 SNN 中, 神经元输出 (即激活) 不是用实值激活编码的, 而是用二元尖峰序列编码的。与传统神经网络相比, 使用 SNN 的动机植根于 SNN

的特殊计算方面，尤其是神经输出激活的高度稀疏性。传统卷积神经网络（CNN）结构成熟，其特点是处理单元（PE）的大型空间阵列在激活稀疏性面前仍然高度未充分利用。我们提出了一种新的架构，该架构针对具有高度激活稀疏性的卷积SNN（CSNN）的处理进行了优化。在我们的体系结构中，主要策略是使用较少但利用率较高的PEs。用于执行卷积的PE阵列仅与内核大小一样大，只要有峰值需要处理，所有PE都可以处于活动状态。通过将特征映射（即激活）压缩到队列中，然后逐峰处理，可以确保这种恒定的峰值流。这种压缩是在运行时使用专用电路执行的，从而实现了自定时调度。这允许处理时间与峰值数量直接成比例。采用一种称为记忆交错的新型记忆组织方案，使用多个小型并行片上ram高效地存储和检索单个神经元的膜电位。每个RAM都通过硬接线连接到其PE，减少了开关电路，并允许RAM位于各自PE附近。我们在FPGA上实现了所提出的架构，与其他实现相比，实现了显著的加速，同时需要更少的硬件资源并保持较低的能耗。

#### 844. 稀疏主动卷积尖峰神经网络的高效硬件加速

Runyu Yu 等，2022. 3. 21，北京邮电大学

简介：针对目前通用的语义文本相似度计算方法难以利用科学学术会议数据的语义信息的问题，提出了一种融合领域特征的科学学术会议语义相似度计算算法。首先，通过实体识别和关键词提取获得会议的领域特征信息，并将其作为特征和会议信息输入伯特网络。暹罗网络的结构用于解决BERT的各向异性问题。将网络的输出合并并归一化，最后使用余弦相似度计算两个会话之间的相似度。实验结果表明，SBFD算法在不同的数据集上都取得了较好的效果，与比较算法相比，斯皮尔曼相关系数有一定的提高。

## 845, 基于故障的数据采集修复脑-机接口

Cailin Winston, 2022. 3. 20,

华盛顿大学

简介：脑-机接口（BCI）对大脑记录的神经信号进行解码和/或用编码的神经信号刺激大脑。BCI 跨越硬件和软件，在恢复医学中有着广泛的应用，从通过假肢和机器人肢体恢复运动，到通过拼写器恢复感觉和交流。BCI 在诊断医学中也有应用，例如为临床医生提供检测癫痫发作、睡眠模式或情绪的数据。

尽管有这样的承诺，但由于可靠性和稳健性方面的挑战，BCI 尚未被长期、日常使用，这是所有场景中安全运行所必需的。确保安全操作目前需要数小时的手动数据收集和重新校准，包括患者和临床医生。然而，数据收集并不是为了消除 BCI 中的特定故障。本文提出了一种新的方法来描述、检测和定位 BCI 中的故障。具体来说，它提出了部分测试预言器作为检测故障的方法，而切片函数作为将故障定位到输入数据或用户执行的相关任务中的特征模式的方法。通过有针对性的数据采集和再培训，该方法提高了 BCI 的正确性。我们在五个 BCI 应用中评估了所提出的方法。结果表明，所提出的方法（1）精确地定位故障，（2）通过基于有针对性的、基于故障的数据采集的再培训，可以显著降低故障频率。这些结果表明，提出的方法是修复故障 BCI 的一个有希望的步骤。

## 846, 基于条件生成对抗网络的驾驶异常检测

YuningQiu 等, 2022. 3. 15

德克萨斯大学达拉斯分校, 本田研究所

内容：异常驾驶检测在高级驾驶员辅助系统（ADAS）中是一个重要的问题。重要

的是尽早识别潜在的危險情况，以避免潜在的事故。本研究提出了一种无监督的方法，使用条件生成对抗网络（GAN）来量化异常驾驶情况。该方法通过之前观察到的信号来调整模型，预测了将要到来的场景。该系统使用预测信号和实际信号之间输出的差值，来作为量化异常程度的指标。

系统采用以驾驶员为中心的方法，收集来自驾驶员的生理信号和车辆上控制器局域网总线（CANBus）的信号。该方法是通过卷积神经网络（CNN）用于提取判别信息特征，并用长短时记忆（LSTM）细胞捕捉时间信息。

这项研究使用异常驾驶数据集（DAD）进行评估，包括 250 小时的手动标注的驾驶事件。实验结果表明带有标注事件的记录更有可能是异常的，比如避免路上行人和违反交通规则，与没有任何事件标注的记录相比，具有更高的异常分数。结果表明异常分数越高的驾驶驱动和其他驾驶驱动相比，更具有危险性，也更不容易在道路上看到，该结果验证本文提出的无监督方法。

#### 847. 基于带偏移量周期性的时空关注度多图卷积网络网约车需求预测

DongXing 等，2022. 3. 25

北京航空航天大学，北海大学

内容：打车服务正在成为城市交通的主导部分。为了提高打车服务的效率，准确预测交通需求是一项根本性的挑战。在本文中，我们从网络结构和数据集公式两个方面来解决这个问题。在网络设计方面，我们提出了一种时空注意多图卷积网络（STA-MGCN）。STA-MGCN 中的时空层通过时间注意机制和时间门卷积捕获时间相关性，通过多图卷积捕获空间相关性。引入特征聚类层来学习潜在的区域函数，减少计算负担。对于数据集公式，我们提出了一种新的方法，该方法考虑了带偏

移的周期性传输特性。不仅使用同一时间段内的历史数据，还包括从昨天到上周的前后相邻时间段内的历史订单需求。在纽约、芝加哥和成都三个真实数据集上进行的大量实验表明，该算法实现了最先进的打车需求预测性能。

#### 848, 跨多个社交媒体平台的可解释错误信息检测

RaheeWalamhe 等, 2022. 3. 23

印度浦那共生技术研究所, 共生国际应用人工智能中心

内容: 因为很大一部分人口依赖互联网获取信息, 所以网络信息处理 (WIP) 对现代社会产生了巨大影响。社交媒体平台提供了传播信息的渠道, 也为传播错误信息、在人群中制造混乱和恐惧提供了温床。检测错误信息的技术之一是基于机器学习的模型, 然而, 由于多个社交媒体平台的可用性, 单独开发和训练基于人工智能的模型已成为一项繁琐的工作, 因此, 采用可解释的人工智能技术至关重要。作者在这篇文章中的贡献为: 开发可跨多个社交媒体平台有效用于错误信息分类的分类器框架 (当可用于训练的数据有限时, 这种方法会很有帮助); 上述架构的 DANN 实现优于 CoAID 数据集 [Cui, 2020] 上的最新结果; 与 COVID-19 疫苗接种相关的新型错误信息数据集 (MiSoVac) 的开发, 从新闻平台和社交媒体网站收集并展示了所提出的方法; 用于解释目标标签的基于 XAI LIME 的方法在已开发的 DANN 模型中调用了更高的可信度, 用于广义的可解释错误信息检测, 值得关注的是, 作者们的开发能够对相似领域的多个数据进行泛化的技术, 在时间、处理和效率方面开发经济方法, 无需在单个平台上训练模型, 也考虑到社交媒体中的领域差异收敛到一般特征, 并伴随着可解释的、值得信赖的采用范式来解决领域适应和可解释性。

**849**，当我们有显著性方法时，为什么要用注意力作为解释？

Jasmijn Bastings 等，2020.11.12

Google Research (期刊)

文章指出，当前许多可解释领域的方法使用注意力机制，但如果使用注意力作为解释的目标是忠实地为输入分配重要性权重，那么为什么要使用注意力而非为实现该目的而设计的众多现有输入显著性方法？

文章首先提出对注意力机制用于可解释性领域的质疑，并结合当前研究者们关于注意力机制争辩的一些工作，讨论了用注意力机制分析问题的合理性。接下来，文章提出了三种显著性方法，即基于梯度的方法 (gradient-based methods)，基于传播的方法 (propagation-based methods)，基于遮挡的方法 (occlusion-based methods)。在介绍完显著性方法后，文章将它们与注意力机制进行对比，指出在可解释领域中，显著性方法要比注意力机制更加忠实可靠。此外，文章肯定了注意力机制在其他领域的应用，并指出了显著性方法仍存在的问题。

**850**，尖峰神经流二进制算术

James B. Aïmone 等，2022.3.23

美国桑迪亚国家实验室

内容：抽象布尔函数和二进制算术运算是标准计算范式的核心。因此，计算领域的许多进展都集中在如何提高这些操作的效率，以及探索它们可以计算什么。为了更好地利用新的计算范例的优点，重要的是考虑它们提供的独特的计算方法。然而，对于任何特殊用途的协处理器，布尔函数和二进制算术运算非常有用，尤

其是通过在设备上预处理和后处理数据来避免不必要的 I/O 打开和关闭协处理器。这尤其适用于尖峰神经形态结构，其中这些基本操作不是基本的低级操作。相反，这些功能需要具体实现。在这里，我们讨论了一种有利的流式二进制编码方法以及一些设计用于精确计算基本布尔运算和二进制运算的电路的含义。

## 851, 稀疏主动卷积尖峰神经网络的高效硬件加速

JanSommer 等,

电气与电子工程师协会

内容: 峰值神经网络 (SNN) 在基于事件的事务中进行计算, 以实现比标准神经网络更有效的计算。在 SNN 中, 神经元输出 (即激活) 不是用实值激活编码的, 而是用二进制尖峰序列编码的。与传统神经网络相比, 使用 SNN 的动机植根于 SNN 的特殊计算方面, 特别是神经输出激活的高度稀疏性。传统卷积神经网络 (CNN) 的成熟架构具有大型处理元素 (PE) 空间阵列, 在面对激活稀集中时, 这些元素仍然高度未得到充分利用。我们提出了一种新颖的架构, 该架构针对卷积 SNN (CSNN) 的处理进行了优化, 具有高度的激活稀疏性。在我们的架构中, 主要策略是使用较少但利用率较高的 PE。用于执行卷积的 PE 数组仅与内核大小一样大, 只要有要处理的峰值, 就允许所有 PE 处于活动状态。通过将功能映射 (即激活) 压缩到队列中来确保这种恒定的峰值流, 然后可以逐个峰值进行处理。这种压缩是在运行时使用专用电路执行的, 从而导致自定时调度。这允许处理时间直接与峰值数量一起扩展。一种称为记忆交错的新型记忆组织方案用于使用多个小型并行片上 RAM 有效地存储和检索单个神经元的膜电位。每个 RAM 都硬连线到其 PE, 从而减少了开关电路, 并允许 RAM 位于相应 PE 附近。我们在 FPGA

上实施了所提出的架构，与其他实现相比，实现了显著的加速，同时需要更少的硬件资源并保持较低的能耗。

### **852，用于物联网流量生成的知识增强 GAN**

ShuodiHui 等，2022. 3. 18

清华大学电子系北京信息科学与技术研究中心，

国家计算机网络技术小组 / 中国协调中心

内容：网络流量数据有助于了解物联网（IoT）行为并提高现实世界中的物联网服务质量。然而，大规模的物联网交通数据很少被访问，而且隐私问题也阻碍了现实的数据共享，即使是匿名的个人身份信息也是如此。研究人员建议生成合成物联网流量，但未能涵盖广泛的现实世界物联网设备提供的多种服务。在这项工作中，我们迈出了通过知识增强生成对抗网络（GAN）框架生成大规模物联网流量的第一步，该框架引入了语义知识（例如，位置和环境信息）和各种网络结构知识。通过知识图的物联网设备。我们使用条件机制来整合物联网流量生成的知识和设备类别。然后，我们采用 LSTM 和自注意力机制来捕获流量序列中的时间相关性。广泛的实验结果表明，我们提出的模型生成的合成物联网流量数据集在数据保真度和应用方面优于最先进的基线。此外，我们提出的模型能够通过仅在知识增强的小型真实数据集上进行训练来生成真实数据。

### **853，关于图像分类的端到端的深度学习架构**

MuhanZhang 等，2022. 3. 25

华盛顿大学

内容：本文提出了一种接受任意结构图的新型神经网络结构，主要解决了图分类问题中的两个挑战：

(1) 如何提取有用的特征来描述那些编码在图中的丰富信息，以达到分类的目的；

(2) 如何以有意义且一致的顺序读取图。

从实际的操作来看，这两个 challenge 分别对应图卷积以及图池化。对此，为了解决第一个挑战，文中设计了一个局部图卷积模型，并给出了它与两个图核的关系；为了解决第二个挑战，设计了一个新的 SortPooling 层，以一致的顺序对图顶点进行排序，这样传统的神经网络就可以在图上进行训练；并通过 top-k 选取对卷积之后的图进行池化。

本文的贡献如下：

(1) 提出了一种新颖的端到端深度学习结构用于图分类。它直接接受图作为输入，不需要任何预处理。

(2) 提出了一种新的空域图卷积层来提取多尺度的顶点特征，并与目前流行的图核进行了类比来解释其工作原理。

(3) 提出了一个新的 SortPooling 层来对顶点特征进行排序，而不是对它们进行汇总，这样可以保留更多的信息，并允许我们更够从全局图的拓扑结构中学习。

(4) 在基准图分类数据集上的实验结果表明，深度图卷积神经网络(DGCNN)与先进的图核相比具有很强的竞争力，并且在图分类方面明显优于许多其他深度学习方法。

**854**，如何让通用人工智能系统能够“识数”

徐英瑾，2021.9.25

复旦大学哲学院

摘要：“识数”的字面意思就是说，数字表达式的使用者需要知道这些表达式日常实践中的具体意义，如知道“十公里”算不算“路远”，“十万元”算不算“昂贵”。主流人工智能系统虽然都是数码化的，但却未必真正“识数”，因为它们都无法以一种灵活的方式来理解数字表达式在特定的日常语境中究竟意味着什么，并在这种理解的基础上进行复杂的语用推理。为了使得人工智能系统能够“识数”，工作思路应是在非公理化推理系统的平台上来构建数字表征。训练非公理化推理系统使之“识数”的指导原则，就是先教会它表征那些最不抽象的数学概念，再一步步进展到更抽象的概念。而这一教学步骤的安排，与我们教导儿童识数的路线图是有点相似的。在这个面向机器的数字教学法中，我们将以对“逻辑专名”的教学作为对数字表征教学的重要前导。

### 855，使用 EEG 的脑机接口进行假肢感觉运动控制的运动预测

Corentin Piozin 等，2022.3.17

法国巴黎中上人研究院巴黎大学综合神经学与认知中心，法国罗伯特·梅勒·德奥比涅研究所

内容：在这项研究中，作者为市售肌电假体的修改版本（Myobock ©，Ottobock）提供基于脑机接口的感觉运动控制。新系统使用用户的脑电图（EEG）信号以及一个手镯产生的振动作为输入。该手镯包含振动电机，其频率与安装在假肢指尖上的力敏电阻器（FSR）测量的力成正比。在对 7 名身体健全者和 4 名截肢者进行的实验中，使用了三种不同特征提取方法（CSP、WD、GSO）的四种组合来构

建两种不同记录系统收集的 EEG 信号的特征向量。三种机器学习算法（人工神经网络、线性核函数的支持向量机和径向基函数的支持向量机）。该研究为使用使用电极较少的 EEG 系统控制肌电假体的运动提供了概念验证。

**856**, 基于多尺度滤波器组的时间、空间特征的集成学习方法用于运动图像脑电图分类

LiangShengZheng 等, 2022. 3. 20

中科院深圳先进技术研究院广东机器人与智能系统重点实验室, 中国科学院大学, 重庆邮电大学

内容: 从脑电图 (EEG) 解码运动意图是基于运动想象的脑机接口 (MI-BCI) 的关键部分。为了帮助神经肌肉残疾的残疾人通过 BCI 恢复运动功能, 需要建立一种高效稳定的分类算法来解码脑电信号中包含的运动意图。然而, EEG 信号是非平稳的, 个体之间差异很大。在这项工作中, 我们提出了一种基于时间、空间特征和多尺度滤波器组的集成学习方法, 称为 TSMFBEL, 旨在为 MI EEG 分类设计具有强泛化能力的集成分类器模型。为了获得多样化的集成分类器, 将原始脑电数据通过自举抽样方法划分为具有样本多样性的子集, 然后通过多尺度滤波器组方法分解为具有时频分布多样性的时频子集。对于每个时频子集, 从时域、空间域和时空域中提取具有域多样性的特征, 并根据每组特征训练具有多样性的异构分类器。为了获得最优决策, 我们将集成策略描述为最小分类误差优化问题, 并提出基于 L2 范数的集成分类器权重优化方法, 最后通过加权融合对集成分类器的决策进行整合。所提出的方法在两个公共数据集 (BCI Competition IV Dataset IIa 和 BCI Competition IV Dataset IIb) 上进行了评估, 并将结果与 最先

进方法的分类方法进行比较。实验结果表明，所提出的 TSMFBEL 算法能够有效地构建多样化的集成分类器，在两个数据集上的平均分类准确率分别为 88.80% 和 86.53%，是目前最先进的方法中最高的，结果的标准差也是最低的。优异的分类性能表明，该算法在脑电信号解码方面具有巨大潜力。

### 857, 基于自适应均值残差损失的人脸年龄稳健估计

Ziyan Zhao 等, 2022. 4. 1

南洋理工大学

内容：自动面部年龄估计在多媒体分析中有着多种实际应用，例如视频监控和人机交互。然而，由于老龄化过程的随机性和模糊性，年龄评估具有挑战性。大多数关于该主题的研究工作都将该任务视为年龄回归、分类和排名问题之一，并且不能很好地利用年龄分布来表示年龄模糊的标签。在这项工作中，我们提出了一个简单而有效的损失函数，用于通过分布学习进行稳健的面部年龄估计，即自适应平均残差损失，其中，平均损失惩罚估计的年龄分布的平均值和地面真实年龄之间的差，而残差损失惩罚分布中动态 top-K 之外的年龄概率熵。数据集 FG-NET 和 CLAP2016 中的实验结果验证了建议损失的有效性。

### 858, 驾驶员碰撞预警的神经符号混合方法

Kyongsik Yun 等, 2022. 4. 1

加州理工学院, 天普大学

论文内容：当前自动驾驶系统有两种主要的算法方法：(1) 端到端系统，其中单个神经网络将学习，如何把感官输入直接映射为适当的警告和驾驶反应。(2)

一种中介混合识别系统，通过组合检测每个语义特征的独立模块来创建一个系统。虽然一些研究人员认为深度学习可以解决任何问题，但另一些研究人员认为，需要一种更加工程化和符号化的方法来处理数据较少的复杂环境。特别地，有时候如果深度学习模型不起作用，调试可能会非常困难，并且深度学习模型对数据分布的变化非常敏感，过度泛化可能会引发问题。通常很难证明它为什么有效或无效。因此本文提出将基于深度学习的对象识别跟踪与自适应神经符号网络代理相结合，称为非公理推理系统（NARS），它可以通过基于感知序列来构建概念，并适应其环境。

非公理推理系统（NARS）是一种智能模型，可以在不同领域进行学习和推理。NARS 是一个可以从现有知识中得出初步结论，并仅在不确定和不完整信息的情况下进行推断的系统。即使没有足够的时间彻底评估所有可能性，NARS 也能迅速做出反应。NARS 不仅可以直接输入人类的逻辑，还可以通过与测试环境交互，从自身的成功和失败经验中学习。在实验的模拟环境中，该方法使用雷达传感器提高目标检测距离的限制，并将基于深度学习的目标检测与神经符号模型相结合，以实现车辆检测功能。基于实验结果，他们所提出的自适应再训练模型，与使用 COCO 数据的预训练模型相比，在对象识别的性能表现上有了较大的提升。与 COCO 数据预训练模型中的 IOU 0.31 相比，本文提出的自适应再训练模型中的 IOU 对象识别性能提高了 0.65。

**859**，领域知识驱动的可解释目标条件交互式弹道预测的伪标签

LingfengSun 等，2022. 3. 30

加州大学伯克利分校，乔治亚理工学院，本田研究院

内容：高度交互场景中的运动预测是自动驾驶中的一个具有挑战性的问题，在这种场景下，需要准确预测交互代理的联合行为，以确保自动驾驶汽车的安全高效导航。在这个领域中，目标条件方法由于其性能优势和捕获轨迹分布中的多模态的能力而受到越来越多的关注。在本文章中，作者们研究了目标条件框架下的联合轨迹预测问题。特别是引入了基于条件变分自动编码器（CVAE）的模型，以将不同的交互模式显式编码到潜在空间中。然而，作者们发现香草模型遭受后塌陷，无法根据需要诱导信息丰富的潜在空间，为了解决这些问题，提出了一种新的方法来避免KL 消失并用伪标签诱导可解释的交互式潜在空间。伪标签允许使用者在交互中加入任意领域知识。在本文章中，作者们使用一个说明性玩具的示例来激发所提出的方法。此外，通过定量和定性评估在 Waymo Open Motion 数据集上验证了文章中提到的框架。

## 860, 部署中的可解释机器学习

U-Bhatt 等, 2020.6.10

卡内基梅隆大学

可解释机器学习旨在通过特征重要性评分、反事实解释和有影响力的样本等技术，为各种利益相关者提供对模型行为的洞察。然而，在这一领域的最新进展中，没有对组织如何在实践中使用这些技术进行调查。本研究探讨了组织如何看待和使用利益相关者消费的可解释性。我们发现，大多数部署不是针对受模型影响的最终用户，而是针对机器学习工程师，他们使用可解释性来调试模型本身。实践中的可解释性与公共透明度的目标之间存在差距，因为解释主要服务于内部利益相

关者，而不是外部利益相关者。我们的研究综合了当前可解释性技术的局限性，这些技术阻碍了终端用户使用它们。为了促进最终用户交互，我们开发了一个框架，用于建立明确的可解释性目标，包括关注规范性需求。

### 861, 用于混合可塑性神经计算的相变记忆传递突触

Syed Ghazi Sarwat 等, 2022. 3. 28

IBM 欧洲研究院

内容：在哺乳动物的神经系统中，各种突触可塑性规则在广泛的时间尺度上单独或协同作用，以促进学习和记忆的形成。因此，在神经拟态计算平台中，非常需要能够忠实表达这种多时间尺度可塑性机制的人工突触。尽管一些可塑性规则已经通过精细的互补金属氧化物半导体和忆阻电路进行了仿真，但缺乏具有可调动态的长期和短期可塑性的设备级硬件实现。在这里，我们介绍了一种相变记忆突触，它利用相位配置的非易失性和场效应调制的易失性来实现可调可塑性。我们表明，这些混合可塑性突触可以启用可塑性规则，例如有助于动态环境建模的短期脉冲时间依赖性可塑性。此外，我们展示了记忆传递突触在实现 Hopfield 神经网络加速器以解决组合优化问题方面的功效。

### 862, 一种用于脉冲神经网络的仿生非线性神经元

Cheng Jin, 2022. 3. 30

麻省理工学院 MIT

内容：脉冲神经网络（SNN）因其处理时间信息的能力和低功耗而引起了研究人员的兴趣。然而，目前最先进的方法限制了它们的生物学合理性和性能，因为它

们的神经元通常是建立在简单的泄漏积分和火灾（LIF）模型上的。由于高度的动态复杂性，现代神经元模型很少在 SNN 实践中实现。在这项研究中，我们采用了相平面分析（PPA）技术，这是一种在神经动力学领域经常使用的技术，来整合一个最近的神经元模型，即 Izhikevich 神经元。根据神经科学进展中的发现，伊兹克维奇神经元模型在生物学上是合理的，同时保持了与 LIF 神经元相当的计算成本。通过利用所采用的 PPA，我们已经完成了将用改进的 Izhikevich 模型构建的神经元应用于 SNN 实践，称为标准化的 Izhikevich 紧张（SIT）神经元。对于性能，我们评估了自建 LIF 和 SIT 组成 SNN 中图像分类任务的建议技术，即静态 MNIST、时尚 MNIST、CIFAR-10 数据集和神经形态 N-MNIST、CIFAR10-DVS 和 DVS128 手势数据集上的混合神经网络（HNN）。实验结果表明，所提出的方法在几乎所有测试数据集上表现出更真实的生物学行为的同时，达到了相当的准确性，证明了这种新策略在弥合神经动力学和 SNN 实践之间的差距方面的有效性。

### 863, 关于图像分类的端到端的深度学习架构

Muhan Zhang, 2022. 3. 30

华盛顿大学

内容：本文提出了一种接受任意结构图的新型神经网络结构，主要解决了图分类问题中的两个挑战：

- (1) 如何提取有用的特征来描述那些编码在图中的丰富信息，以达到分类的目的；
- (2) 如何以有意义且一致的顺序读取图。

从实际的操作来看，这两个 challenge 分别对应图卷积以及图池化。对此，为了

解决第一个挑战，文中设计了一个局部图卷积模型，并给出了它与两个图核的关系；为了解决第二个挑战，设计了一个新的 SortPooling 层，以一致的顺序对图顶点进行排序，这样传统的神经网络就可以在图上进行训练；并通过 top-k 选取对卷积之后的图进行池化。

本文的贡献如下：

(1) 提出了一种新颖的端到端深度学习结构用于图分类。它直接接受图作为输入，不需要任何预处理。

(2) 提出了一种新的空域图卷积层来提取多尺度的顶点特征，并与目前流行的图核进行了类比来解释其工作原理。

(3) 提出了一个新的 SortPooling 层来对顶点特征进行排序，而不是对它们进行汇总，这样可以保留更多的信息，并允许我们更够从全局图的拓扑结构中学习。

(4) 在基准图分类数据集上的实验结果表明，深度图卷积神经网络(DGCNN)与先进的图核相比具有很强的竞争力，并且在图分类方面明显优于许多其他深度学习方法。

#### 864, 用于知识图谱中链接预测的复杂图卷积网络

AdnanZeb 等, 2022. 3. 26

厦门大学信息学院, 巴基斯坦 COMSATS 大学, 马拉坎德大学, 澳门大学

内容: 知识图谱 (KG) 嵌入模型将节点和边映射到固定长度的向量, 并获得节点的相似度作为评分函数的输出, 以预测节点之间的缺失链接。基于图卷积网络 (GCN) 的 KG 嵌入方法最近受到了广泛关注, 因为它们能够将相邻节点的信息添加到节点的嵌入中。然而, 现有的 GCN 主要基于具有高失真的实值嵌入, 尤其

是在对具有不同几何结构的图进行建模时。在本文中，我们提出了复杂图卷积网络 (ComplexGCN)，它是标准 GCN 在复杂空间中的一种新扩展，将复杂几何的表达能力与 GCN 相结合，以提高 KG 组件的表示质量。所提出的 ComplexGCN 包括一组复杂的图卷积层和一个基于 PARATUCK2 分解的复杂评分函数：前者将相邻节点的信息包含在节点的嵌入中，而后者利用这些嵌入来预测节点之间的新链接。与两个最近的标准链路预测数据集上的现有方法相比，所提出的模型展示了增强的性能。

## 865, 深度“汉字屋”与通用人工智能两难

颜青山, 2018.9

华东师范大学哲学系

摘要：更精确地重述“汉字屋”论证的条件，“汉字屋”中的塞尔以不同于中国人的方式“学会”汉语语法，他从而试图反驳强人工智能的论证是失败的，尽管其结论依然可以成立。心灵智能的奠基性特征恰好是基于语法或语形的实践理解，完全的实践理解将通过神经肌肉系统重塑大脑，并且具有部分可逆性。完全的实践理解需要周期性的训练维持，这一特征将导致通用人工智能的两难：如果通用人工智能是心灵智能，那么它不可能全面超越人类智能；如果通用人工智能全面超越人类智能，那么它不可能是心灵智能。

为什么要参照人类智能来看人工智能呢？这是因为智能的特征问题是一个开放的经验问题，我们只能参照已有心灵的核心特征来考虑人工智能的心灵可能性。也许，有人会反驳说，人类智能可以全面超越黑猩猩的智能，那么是否可以设想通用的强人工智能也可以在各个方面超越人类呢？答案当然是否定的，因为人类智

能并没有全面超越黑猩猩，例如在认知单元上人类恰恰弱于黑猩猩。最后，既然超级智能不可能是人类的智能，如果它出现了，我们可以将其仅仅看作手段。这将告慰那些过度担忧伦理问题的文人们。至于超级智能作为工具可能引起的其他伦理问题，我们完全可以参照对待其他工具的方式处理。

## 866， 机器也需教育？论通用人工智能与教育学的革新

刘凯、胡祥恩、王培，2018.2

华中师范大学心理学院、孟菲斯大学智能研究院、天普大学计算机与信息科学系  
摘要：历史悠久且底蕴深厚的教育学如今却处于被边缘化的尴尬境地，通用人工智能的兴起为教育学复兴带来契机。总体而言，人工智能分为专用(弱)人工智能和通用(强)人工智能两大分支，二者技术路线完全不同。专用人工智能侧重对“智能”行为外在的模拟，通用人工智能则致力实现机器内在的思维和情感，认为智能是一个系统在知识和资源相对不足情况下的适应能力，强调自身经验对个体塑造的决定性作用。在教育领域中，通用人工智能不仅破除了羁绊专用人工智能的技术崇拜与人文关怀迷失两大问题，更表现了与人类教育活动极相似的特征：首先是“教”，人类学习理论对通用人工智能系统的学习过程同样有效；其次是“育”，通用人工智能系统在学习时也同样需要借助经验积累来实现育化的效果。因此，在通用人工智能的启迪下，教育学将打破“人”的先天束缚，在人类教育理论与实践的基础上吸纳和统合机器教育，在更宽广的“人-机”二元主体视角下探究教育与学习的一般性规律，向“大学科”和“大科学”的方向迈进。

## 867， 基于 CNN 的二维人类情绪与多目标进化算法选择 EEG 通道的区分（脑机接

口)

Luis Alfredo Moctezuma 等, 2022. 3. 3

挪威科技大学, 国际综合睡眠医学研究所, 筑波大学

内容: 在这项研究中, 我们探讨了不同水平的情绪强度(Arousal)和愉快(Valence)如何反映在脑电图(EEG)信号中。我们对来自 DEAP 公共数据集的 32 名受试者的 EEG 数据进行了实验, 其中受试者使用 60 秒视频刺激以引发不同水平的唤醒/效价, 然后使用自我评估从 1 到 9 自我报告评分人体模型 (SAM)。EEG 数据经过预处理并用作卷积神经网络 (CNN) 的输入。首先, 32 个 EEG 通道用于计算每个受试者可获得的最大准确度水平, 以及使用来自所有受试者的数据创建单个模型。一次使用一个通道重复实验, 查看特定渠道是否包含更多信息以区分低唤醒/高唤醒/效价。结果表明, 与使用所有 32 个通道相比, 使用一个通道的准确度较低。然后使用非支配排序遗传算法 II (NSGA-II) 设计 EEG 通道选择的优化过程, 目的是获得具有高精度识别的最佳通道组合。遗传算法使用所有 32 个通道的染色体表示来评估所有可能的组合, 并迭代测试来自不同群体中每个染色体的 EEG 数据, 以解决两个无约束目标; 最大限度地提高分类精度并减少分类过程所需的脑电图通道数量。从 Pareto 前沿获得的最佳组合表明, 只有 8-10 个通道就可以满足这一条件, 并为更轻量级的 EEG 系统设计以进行情绪识别奠定了基础。在最好的情况下, 结果显示精度高达 1.00 使用八个 EEG 通道的低唤醒与高唤醒, 以及仅使用两个 EEG 通道的低与高价为 1.00。这些结果对于需要使用可穿戴 EEG 进行自动情绪识别的研究和医疗保健应用来说是令人鼓舞的。

868, EEG-ITNet: 用于运动图像分类的可解释初始时间卷积网络 (脑机接口)

Abbas Salami 等, 2022. 3. 22

埃塞克斯大学计算机科学与电子工程学院、心理学院

内容: 近年来, 神经网络, 尤其是深度架构在脑机接口 (BCI) 领域的脑电信号分析方面受到了广泛关注。在这个正在进行的研究领域中, 端到端模型比需要信号转换预分类的传统方法更受青睐。它们可以消除对专家先验信息和手工特征提取的需求。然而, 尽管文献中已经提出了几种深度学习算法, 在对运动或心理的分类任务上取得了很高精度, 但它们往往缺乏可解释性, 因此不太受到神经科学界的青睐。这是因为模型参数庞大, 深度神经网络对微小但不相关的判别特征十分敏感。我们提出了一种名为 EEG-ITNet 的端到端深度学习架构, 以及一种更易于理解的方法来可视化网络学习模式。使用初始模块和扩展因果卷积, 我们的模型可以从多通道 EEG 信号中提取丰富的频谱、空间和时间信息, 其复杂性 (就可训练参数的数量而言) 比其他现有的端到端架构 (如 EEG-Inception 和 EEG-TCNet 更低。通过对来自 BCI 竞赛 IV 和 OpenBMI 运动图像数据集的数据集 2a 的详尽评估, EEG-ITNet 与竞争对手相比分类准确率提高了 5.9%, 在统计学上具有显著差异。

### 869, 基于深度重要性抽样的自行车加速评估

(机器学习, 高维不确定性测试)

MansurArief 等, 2022. 4. 4

卡内基梅隆大学, 哥伦比亚大学, 同济大学, 伊利伊诺大学

论文内容: 人工智能可以高效处理大量高维输入数据, 为下游任务提供预测, 提升自动驾驶的性能表现, 但也需要相应的评估手段, 来确保这些复杂的人工智能

驱动技术，在部署时始终可靠安全地运行。在自然环境下，高精度评估自动驾驶汽车（AV）及其复杂子系统的性能仍然是一个挑战，尤其是在故障或危险案例很少的情况下。稀有性不仅需要巨大的样本量才能实现高置信度估计，而且还会导致对真实故障率的危险低估，而且极难检测出来。同时，在正确性保证方面上最先进方法只能计算特定条件下的故障率上限，而这可能会限制其实际应用。

在这项工作中，本文提出了深度重要性抽样（Deep IS）框架，该框架利用深度神经网络获得与最新技术水平相当的有效重要性采样，并且能够将所需样本量减少 43 倍，比原始抽样方法小，以实现 10% 的相对误差，同时产生保守得多的估计。本文在高维实验中估计了其中一个最先进的交通标志分类器的误分类率，实验结果进一步表明即使检测目标非常小，仍然能够保持高效地分类，对比标准方法实现了 600 多倍的效率提升。这突出了 Deep IS 在提供精确估计方面的潜力，即使在高维不确定性的情况下也能提供较高的准确率。未来，作者还计划将上限估计器与深度重要性采样估计器并行使用，以设计更有效的采样停止标准，结合两种估计器的优点，进一步提高智能交通系统可认证评估方法的效率。

## 870. 对 1 型糖尿病患者更安全的血糖控制

（深度学习，离线强化学习）

Harry Emerson 等，2022. 4. 8

Bristol 大学，Southampton 大学医院

内容：混合闭环系统代表了 1 型糖尿病（T1D）患者护理的未来。这些设备通常使用简单的控制算法来选择最佳胰岛素剂量，以将血糖水平维持在健康范围内。在线强化学习（RL）已被用作进一步加强这些设备中葡萄糖控制的方法。与经典

控制算法相比，以前的方法已被证明可以降低患者风险，并改善在目标范围内花费的时间，但在学习过程中容易出现不稳定，通常会导致选择不安全的行为。这项工作对离线 RL 进行了评估，将其作为制定临床有效剂量政策的一种手段，而无需与患者互动。本文研究了 BCQ、CQL 和 TD3-BC 在 UVA Padova 葡萄糖动力学模拟器中管理九名虚拟患者血糖的效用。当在线 RL 方法所需的数据不足十分之一时，这项研究表明，与最先进的基线相比，离线 RL 可以显著增加健康血糖范围内的时间。这是在没有任何相关的低血糖事件增加的情况下实现的。离线 RL 还可以纠正常见且具有挑战性的情况，如不正确的药丸剂量、不规则的用餐时间和次优的训练数据。

## 871, 可解释人工智能研究综述（迈向医学可解释人工智能）

Erico Tjoa 等, 2022. 4. 8

IEEE

最近，从图像处理到自然语言处理，人工智能和机器学习在许多任务中都表现出非凡的表现，尤其是随着深度学习的出现。随着研究的进步，他们已经涉足了许多不同的领域和学科。其中一些需要高度的问责制和透明度，例如医疗部门。因此，需要对机器决策和预测进行解释以证明其可靠性。这需要更高的可解释性，这通常意味着我们需要了解算法背后的机制。不幸的是，深度学习的黑盒性质仍未解决，许多机器决策仍然知之甚少。我们对不同研究工作提出的可解释性进行了回顾，并对它们进行了分类。不同的类别在可解释性研究中表现出不同的维度，从提供“明显”可解释信息的方法到复杂模式的研究。通过将相同的分类应用于医学研究中的可解释性，希望（1）临床医生和从业人员随后可以谨慎使用这些

方法，(2) 对可解释性的见解将在医学实践中产生更多考虑，以及 (3) 倡议鼓励推进以数据为基础、以数学和技术为基础的医学教育。

## 872, 一种通过可解释 AI 系统指导图像恢复对比方法

Aboli Marathe 等, 2022. 4. 6

印度大学, 普纳计算技术学院, 印度共生技术研究所

内容: 自动驾驶汽车和无人机等现代应用在很大程度上依赖于强大的物体检测技术。然而, 天气情况会阻碍物体的可检测性, 并对它们的导航和可靠性构成严重威胁。因此, 需要有效的去噪、去雨和恢复技术。生成对抗网络和转换器已被广泛用于图像恢复。然而, 这些方法的训练往往不稳定且耗时。

此外, 当用于对象检测 (OD) 时, 尽管图像清晰, 但这些方法生成的输出图像可能会提供不令人满意的结果。在这项工作中, 作者们提出了一种对比方法来缓解这个问题, 通过评估恢复模型在训练期间和训练后生成的图像。这种方法利用 OD 分数与注意力图相结合来预测恢复图像对 OD 任务的有用性。作者们使用条件 GAN 的两个新用例和两种转换器方法进行实验, 以探索所提出的方法对 OD 任务中多天气损坏的鲁棒性。

作者们的方法在沙尘龙卷风和降雪等恶劣天气条件下, 输入图像和恢复图像之间的 mAP 平均增加了 178 倍。文章报告了一些独特的案例, 即更大的去噪不会提高 OD 性能, 相反, 生成的噪声图像显示出良好的结果。文章得出结论, 需要可解释性框架来弥合人类和机器感知之间的差距, 特别是在自动驾驶汽车的鲁棒目标检测方面。

### 873, 基于天然有机材料蜂蜜的忆阻性突触装置

(用于再可生物降解的神经拟态系统中激发神经网络)

Brandon Sueoka 等, 2022. 4. 8

华盛顿州立大学

内容: 在未来的神经拟态结构中, 尖峰神经网络 (SNN) 要求硬件设备不仅能够模拟生物突触的基本功能, 如尖峰时间依赖性可塑性 (STDP) 和尖峰速率依赖性可塑性 (SRDP), 而且能够生物降解, 以应对当前电子废物的生态挑战。在不同的设备技术和材料中, 基于天然有机材料的忆阻性突触设备已成为满足这些需求的理想选择。金属-绝缘体金属结构类似于生物突触, 具有低功耗、快速开关速度和突触可塑性模拟, 而天然有机材料具有水溶性、可再生性和环境友好性。在这项研究中, 一种天然有机材料蜂蜜基忆阻器用于 SNN 的潜力得到了证实。该器件表现出自由双极电阻开关、100ns 设置时间和 500ns 重置时间的高开关速度、STDP 和 SRDP 学习行为以及溶解在水中。提出了 STDP 和 SRDP 的直观传导模型。这些结果证明, 蜂蜜基记忆性突触器件有望在绿色电子和可生物降解的神经拟态系统中实现 SNN。

### 874, 脉冲神经网络时空模式传播的解析研究

Ning Hua 等, 2022. 4. 7

复旦大学

内容: 基于矩闭包方法, 构造高斯随机场来定量和解析地描述随机场的动力学特性。该方法为我们研究前馈或递归尖峰神经网络中的同步尖峰传播提供了理论工具。我们发现, 突触后兴奋和抑制电位之间的平衡是合成火链发生所必需的。特

别是在平衡网络中，可以观察到入侵和毁灭的临界数据包大小。我们还推导了异步环境中同步传播的一个充分分析条件，进一步揭示了空间突触结构维持稳定的突触链的可能性。我们的发现与模拟结果非常一致，并帮助我们理解随机点飞行中时空模式的传播。

### **875, 用于跨模态检索的细粒度双向注意力生成和知识辅助网络**

JianWei Zhu 等, 2022. 4. 1

广西师范大学, 西北师范大学

内容: 通常, 大多数现有的跨模态检索方法只考虑全局或局部语义嵌入, 缺乏对象之间的细粒度依赖关系。同时, 通常会忽略模态之间的相互转换也有利于模态的嵌入。鉴于这些问题, 我们提出了一种称为 BiKA (双向知识辅助嵌入和基于注意力的生成) 的方法。该模型使用双向图卷积神经网络来建立对象之间的依赖关系。此外, 它采用双向基于注意力的生成网络来实现模态之间的相互转换。具体来说, 知识图谱用于局部匹配以约束模态的局部表达, 其中生成网络用于相互转换以约束模态的全局表达。此外, 我们还提出了一种新的位置关系嵌入网络来嵌入对象之间的位置关系信息。在两个公共数据集上的实验表明, 与许多最先进的模型相比, 我们的方法的性能得到了显著提高。

### **876, 基于自注意力网络的动态图表示学习**

Sankar Aralind, 2022. 4. 6

伊利诺伊大学巴纳香槟分校

内容: 学习图中节点的潜在表示是一项重要且无处不在的任务, 具有广泛的应用,

例如链接预测、节点分类和图可视化。以前的图表示学习方法主要集中在静态图上，然而，许多现实世界的图是动态的并且随着时间的推移而发展。在本文中提出了动态自注意力网络 (DySAT)，这是一种新颖的神经架构，可在动态图上运行并学习捕获结构特性和时间演化模式的节点表示。具体来说，DySAT 通过沿两个维度联合使用自注意力层来计算节点表示：结构邻域和时间动态。文章对两类图进行链接预测实验：通信网络和二分评价网络。论文中的实验结果表明，DySAT 在几个不同的最先进图嵌入 baseline 上具有显著的性能提升。

提出的动态的自注意力网络是一种新型的神经架构，它操作在动态图上，并学习节点表示，以捕捉结构特性和时间演化模式，具有重要的参考意义。

## 877, 人类教学原则能用于通用人工智能吗？

(像教育人一样教育机器)

刘凯等, 2022.4

渤海大学, 华中师范大学, 美国天普大学

摘要：通用人工智能旨在探索智能的一般规律，使其软件系统具有感知、记忆、情感、推理与决策等多种 认知功能，乃至拥有与人类学习者相似的学习能力。本文将通用人工智能系统作为机器学习者，考察人类教学原则对机器学习者的教育干预效果，探究机器教育的可能性。受控实验结果显示，教学时长、教学内容及教学方法的变化都能引发通用人工智能系统的记忆网络结构显著改变。这不仅证实了人类教学原则迁移至通用人工智能系统的可能性，也折射出存在贯通“人一机”的普适的、可证伪的教育客观规律。在人工智能等新兴技术赋能教育的时代背景下，教育学的原生创新仍相对乏力。通用人工智能系统的出现不仅丰富了

教育学研究对象，也拓宽了教育学研究边界。在通用人工智能框架下，认知机器不但能够而且也应该受到教育，用人类教育手段反哺人工智能将成为教育科学今后的重要研究课题之一，教育科学也有望成为促进科学发展的新动力。

### 878, 基于神经流形的动作电位神经网络增强皮质内脑机接口数据

ShangjieZheng 等, 2022. 3. 31

中国科学院大学, 清华大学深圳国际研究院

内容: 脑机接口将大脑中的神经信号转换为控制外部设备的指令。然而, 很难获得足够的训练数据。随着先进机器学习方法的出现, 脑机接口的能力得到了前所未有的提高, 但这些方法需要大量的数据进行训练, 因此需要对有限的可用数据进行数据增强。在这里, 我们使用动作电位神经网络(SNN)作为数据生成器。它被被认为是面向通用人工智能的算法之一, 因为它借用了生物神经元的神经信息处理过程。我们使用 SNN 生成生物可解释并符合原始神经数据的固有模式的神经脉冲信息。实验表明, 该模型可以直接合成新的脉冲序列, 从而提高了 BCI 译码器的泛化能力。动作电位神经模型的输入和输出都是动作电位信息, 这是一种未来可以更好地与 BCI 集成的方法。

### 879, 人类新生儿脑电数据的伪影去除管道

Veluprabhakar Kumaravel 等, 2022. 1. 24

意大利布鲁诺-凯斯勒基金会, 特伦托大学心理 / 脑科学中心

内容: 脑电图 (EEG) 正在成为一种有价值的方法来研究出生后不久的神经认知功能。然而, 从人类新生儿记录中获取高质量的脑电图数据具有挑战性。与成人

和较大婴儿相比，由于新生儿的注意力有限，数据集通常要短得多，而由于主要由无法控制的运动引起的非刻板伪影，数据集通常要短得多。我们提出了新生儿脑电图伪影去除 (NEAR)，这是一种专门为人类新生儿设计的脑电图伪影去除管道。NEAR 基于两个关键步骤：1) 基于局部异常因子 (LOF) 的新型坏通道检测工具，一种鲁棒的异常值检测算法；2) 用于适应新生儿脑电图数据的参数校准程序，算法伪影子空间重建 (ASR)，为移动成人脑电图中的伪影去除而开发。对模拟数据的测试表明，NEAR 在去除具有代表性的新生儿非定型伪影方面优于现有方法。NEAR 在使用两种不同的实验设计 (频率标记和 ERP) 记录的两个发育人群 (新生儿和 9 个月大的婴儿) 上进行了验证。结果表明，NEAR 伪影去除成功地从嘈杂的数据集中重现了已建立的 EEG 响应，具有比现有伪影去除方法获得的更高的统计意义。基于 EEGLA 的 NEAR 管道可在以下位置

## 880, 忆阻器的三组分漂移扩散模型

Clément Jourdana 等, 2022. 4. 7,

维也纳理工大学, 让·昆茨曼实验室

在有界区域内，采用混合 Dirichlet-Neumann 边界条件，分析了半导体中电子、空穴和含氧空位密度的漂移扩散方程组与电势的泊松方程组。该系统描述了忆阻器器件中电荷载流子的动力学。忆阻器可以看作是具有记忆功能的非线性电阻，模仿生物突触的电导反应。在快弛豫极限下，系统的含氧空位密度和电势降低为漂移扩散系统，这在神经形态应用中经常使用。证明了以下结果：全系统在任意维上弱解的整体存在性；全系统解的一致时间有界性和二维快速松弛极限；约化系统的整体存在性和弱强唯一性分析。在一维空间中的数值实验说明了解的行为，

并再现了电流-电压特性中的滞后效应。

### 881, 基于脑电的脑机接口的生理信息数据扩充

Oleksandr Zlatov 等, 2022. 3. 27

德国柏林工业大学

大多数基于 EEG 的脑机接口 (BCI) 需要大量的训练数据来校准分类模型, 这是因为 EEG 数据的高度可变性, 这表现在参与者之间, 但也表现在参与者之间的会话之间 (当然, 从试验到试验)。一般来说, 模型越复杂, 需要的训练数据就越多。我们提出了一种通过从手头的数据集生成新数据来增加训练数据的新技术。与现有技术不同, 我们的方法使用反向和正向投影, 使用源定位和头部模型来修改模型的当前源偶极子, 从而以生理意义上的方式产生参与者间的变异性。在这篇手稿中, 我们解释了该方法, 并展示了参与者独立运动想象分类的初步结果。当使用深度神经网络、浅层神经网络和 LDA 时, 使用所提出的数据增强方法的准确度分别提高了 13、6 和 2 个百分点。

### 882, 神经网络在类人语音识别中的成功与关键失败

Federico Adolfi 等, 2022. 4. 6

(英) 布里斯托大学, (美) 纽约大学

自然听觉和人工听觉原则上可以为给定的问题演化出不同的解决方案。然而, 这项任务的限制可以推动认知科学和听觉工程在质量上趋同, 这表明更密切的相互检查将改善人工听觉系统以及大脑和大脑的过程模型。语音识别——这是一个成熟的探索领域——在人类身上对各种光谱时间粒度的数字变换具有固有的鲁棒

性。高性能神经网络系统在多大程度上解释了这些鲁棒性特征？我们将语音识别实验集中在一个单一的合成框架下，以评估最先进的神经网络作为刺激可计算、优化的观察者。在一系列实验中，我们（1）阐明了文献中有影响力的语音操作之间以及与自然语音之间的关系，（2）展示了机器表现出分布外鲁棒性的粒度，再现了人类的经典感知现象，（3）确定了人类表现的模型预测不同的特定条件，（4）证明了所有人工系统在感知上无法恢复到人类的状态，这是理论和模型构建的关键规范。这些发现鼓励了认知科学和听觉工程之间更紧密的协同作用。

### 883, 基于语义深度云映射和多智能体的全端到端自动驾驶

Oskar Natan 等, 2022.04.12

日本丰桥技术科学大学

针对自动驾驶车辆的点到点导航任务，该论文提出了一种新的深度学习模型，该模型采用端到端和多任务学习方式训练，同时执行感知和控制任务。该模型通过遵循全局规划器制定的路线，来驱使车辆安全自动驾驶。该模型的感知部分，用于对 RGBD 摄像机提供的高维观测数据进行编码，同时执行语义分割，使用语义深度云（SDC）映射，以及交通灯状态和停车标志的预测。然后，控制部分对 GPS 和里程表提供的附加信息进行编解码，来预测潜在特征空间的行车路线。此外，使用两个智能体来处理这些输出，并制定控制策略，以确定转向、油门和制动的级别，作为车辆的最终操作。

该模型在 CARLA 模拟器上进行评估，使用由正常对抗情况和不同天气组成的各种场景来模拟真实世界的条件。此外，论文作者还与一些最新的模型进行了对比研究，调整改善了多个方面的性能。论文作者还对 SDC 映射和多智能体进行了消融

研究，以了解它们的作用和行为。因此，在参数和计算量较少的情况下，论文提出的模型也能获得最高的驾驶分数。

对于未来的研究，作者认为可以在多个视图中使用 RGBD 摄像头，以捕捉车辆附近更多信息。通过更好地理解场景，可以使得模型具有更好的驾驶性能。同时可以考虑，通过在感知模块上添加深度估计解码器来替代深度映射输入。然而，每个像素的估计误差需要仔细考虑，因为这会影响语义深度云映射。

#### **884, YOLO-POSE: 基于对象关键点相似度损失的多人姿态估计算法**

Debapriya Maji 等, 2022. 04. 15

介绍了一种新的无热图的关节检测方法 YOLO pose, 以及基于流行的 YOLO 目标检测框架的二维多人姿态估计。现有的基于热图的两阶段方法是次优的, 因为它们不是端到端的可训练方法, 训练依赖于替代 L1 损失, 这不等同于最大化评估指标, 即对象关键点相似性 (OKS)。我们的框架允许我们端到端地训练模型, 并优化 OKS 度量本身。该模型学习在一次向前传球中联合检测多人的边界框及其相应的 2D 姿势, 从而实现自上而下和自下而上方法的最佳结合。所提出的方法不需要自底向上方法的后处理来将检测到的关键点分组到骨架中, 因为每个边界框都有一个相关的姿势, 从而产生关键点的固有分组。与自上而下的方法不同, 由于所有人都在一次推理中随姿势定位, 因此取消了多个向前传球。YOLO pose 在 COCO 验证 (90.2%AP50) 和测试开发集 (90.3%AP50) 方面取得了最新的成果, 在一次向前传球中超越了所有现有的自下而上方法, 无需翻转测试、多尺度测试或任何其他测试时间增加。与使用翻转测试和多尺度测试来提高性能的传统方法不同, 本文中报告的所有实验和结果都没有增加任何测试时间

## 885, 可解释机器学习中无基础真理的解释性评价

Fan Yang 等, 2022 年 4 月 15 日

美国得克萨斯农工大学

可解释的机器学习 (IML) 在许多现实世界的应用中变得越来越重要, 例如自动驾驶汽车和医疗诊断, 在这些应用中, 解释非常受欢迎, 以帮助人们更好地理解机器学习系统的工作原理并进一步增强他们对系统的信任。然而, 由于场景的多样化和解释的主观性, 我们很少有在 IML 中对生成的解释质量进行基准评估的基本事实。具有解释质量感不仅对评估系统边界很重要, 而且有助于在实际环境中实现对人类用户的真正好处。为了对 IML 中的评估进行基准测试, 在本文中, 我们严格定义了评估解释的问题, 并系统地回顾了现有技术的现有努力。具体来说, 我们用正式的定义总结了解释的三个一般方面(即普遍性、保真度和说服力), 并分别回顾了在不同任务下它们各自的代表性方法。此外, 根据开发人员和最终用户的分层需求设计了统一的评估框架, 可以在实践中轻松采用不同的场景。最后, 讨论了开放问题, 并提出了当前评估技术的一些局限性, 以供未来探索。

## 886, 面向实体肺结节诊断的可靠可解释人工智能模型

Chenglong Wang, 等, 2022. 04. 12

华东师范大学, 南京医科大学

肺癌是世界上死亡率最高的致命癌症。早期发现对肺癌的治疗至关重要。计算机辅助诊断 (CAD) 系统已被开发用于辅助放射科医师进行结节检测和诊断, 大大减轻了医生们的工作量, 同时提高了诊断准确性, 然而, 缺乏模型可靠性和可解释性仍然是其大规模临床应用的主要障碍。在本文中, 作者们提出了一种用于肺

结节诊断的多任务可解释深度学习模型。作者们的神经模型不仅可以预测病变恶性程度，还可以识别相关表现。此外，每个表现的位置也可以可视化以进行视觉解释；作者们提出的神经模型在 LIDC 公共数据集上实现了 0.992 的测试 AUC，在作者们的内部数据集上实现了 0.923 的测试 AUC。此外，作者们的实验结果证明，通过将表现识别任务纳入多任务模型，还可以提高恶性肿瘤分类的准确性，这种多任务可解释模型可以提供一种方案，以便在临床环境中与放射科医生更好地互动。在 LIDC 公共数据集上为 992，在作者们的内部数据集上测试 AUC 为 0.923。此外，作者们的实验结果证明，通过将表现识别任务纳入多任务模型，还可以提高恶性肿瘤分类的准确性，这种多任务可解释模型可以提供一种方案，以便在临床环境中与放射科医生更好地互动。

## 887, 一种实现强化学习的尖峰神经网络结构

Mikhail Kiselev 2022.4.9

俄罗斯楚瓦什国立农学院

目前，尽管提出了大量的 SNN 学习算法，但在脉冲神经网络 (SNN) 中实现学习机制并不能被视为已解决的科学问题。对于强化学习 (RL) 的 SNN 实现也是如此，而 RL 对于 SNN 尤其重要，因为它与从 SNN 应用的角度来看最有希望的领域（例如机器人技术）密切相关。在本文中，我描述了一种 SNN 结构，它似乎可以用于广泛的 RL 任务。我的方法的显着特点是仅使用所有相关信号的尖峰形式——感觉输入流、发送到执行器的输出信号和奖励/惩罚信号。除此之外，在选择神经元/可塑性模型时，我的要求是它们应该可以在现代神经芯片上轻松实现。论文中考虑的 SNN 结构包括由 LIFAT（具有自适应阈值的泄漏集成和激发

神经元)模型的泛化描述的脉冲神经元和一个简单的脉冲时间依赖突触可塑性模型(多巴胺调制可塑性的泛化)。我的概念基于关于 RL 任务特征的非常一般的假设,并且对其适用性没有明显的限制。为了测试它,我选择了一个简单但不平凡的任务,即训练网络以在模拟 DVS 摄像机的视野中保持一个混乱移动的光点。所描述的 SNN 成功解决了这个 RL 问题可以被视为有利于我的方法效率的证据。我选择了一个简单但不平凡的任务,即训练网络以在模拟 DVS 摄像机的视野中保持混乱移动的光点。所描述的 SNN 成功解决了这个 RL 问题可以被视为有利于我的方法效率的证据。我选择了一个简单但不平凡的任务,即训练网络以在模拟 DVS 摄像机的视野中保持混乱移动的光点。所描述的 SNN 成功解决了这个 RL 问题可以被视为有利于我的方法效率的证据。

## 888, 尖峰神经网络的最新进展和新前沿

Duzhen Zhang 等, 2022.03.22

中国科学院大学人工智能学院

近年来,尖峰神经网络(SNN)由于其丰富的时空动力学,各种编码方案和自然适合神经形态硬件的事件驱动特征,在大脑启发智能领域受到广泛关注。随着 SNN 的发展,脑启发智能这个以脑科学成果为灵感、以人工通用智能为目标的新兴研究领域,正变得炙手可热。本文综述了 SNN 的最新进展,并从四大研究课题中探讨了 SNN 的新前沿,包括本质要素(即峰值神经元模型、编码方法和拓扑结构)、数据集、优化算法以及软硬件框架。我们希望我们的调查可以帮助研究人员更好地了解 SNN,并激发新的工作来推进这一领域。

## 889, 可池化分层图表示学习

论文作者: Zhitao Ying, 2022. 03. 10

美国斯坦福大学、加拿大麦吉尔大学、南加州大学

图神经网络 (GNN) 通过有效学习节点嵌入, 彻底改变了图表示学习领域, 并在节点分类和链接预测等任务中取得了最先进的成果。然而, 当前的 GNN 方法本质上是扁平的, 并且不学习图的层次表示——这一限制对于图分类任务尤其成问题, 其目标是预测与整个图关联的标签。本文提出了 DiffPool, 这是一个可微的图池化模块, 它可以生成图的层次表示, 并且可以以端到端的方式与各种图神经网络架构相结合。DiffPool 为深度 GNN 的每一层的节点学习可微分软集群分配, 将节点映射到一组集群, 然后形成下一个 GNN 层的粗化输入。我们的实验结果表明, 与所有现有的池化方法相比, 将现有的 GNN 方法与 DiffPool 相结合, 在图分类基准上的准确率平均提高了 5-10%, 在五分之四的基准数据集上实现了新的最新技术。

## 890, 聚合知识感知图神经网络和自适应关系注意以进行推荐

Yihao Zhang, 2022. 4. 7

重庆理工大学

目前, 基于知识图谱 (KG) 的推荐系统会考虑项目的各个方面来提供准确的推荐。许多研究表明, 利用知识图谱丰富的语义可以有效地提高推荐性能, 并且可以解决数据稀疏性并提供可解释的推荐。然而, 大多数现有的基于 KG 的推荐系统忽略了用户对项目的各种关系赋予不同程度的重要性这一事实。为了解决这个问题, 我们提出了一种基于自适应关系注意 (KGARA) 的知识图谱推荐模型, 该模型可

以捕捉各种用户对不同项目关系的注意。具体来说，我们引入关系嵌入来对 KG 的语义信息进行建模，并通过注意力机制捕捉用户对目标项目每个关系的注意力。此外，我们引入感受域来选择 KG 中目标节点的邻居节点，这大大减轻了计算负担。对三个真实世界数据集的广泛实验表明，所提出的算法比其他最先进的算法有显著改进

## 891，脑工终结时代的来临：通用人工智能机器生产工艺学批判

刘方喜 2022 年 3 月

中国社会科学院文学研究所

能量高度自动化的机器曾经终结手工时代，通用人工智能将成为高度发达的社会大脑所生成的高度自动化的社会机械通用智能，脑工终结时代正在来临。人类有望彻底征服自然力，使物质和精神生产力从人身生物性限制下全面解放出来，在物种关系方面超越人类与自然关系的动物性；扬弃资本之后，人与人社会关系中过度竞争的动物性也将被超越，每个人的手工、脑工劳动将从资本支配下的竞争性、商品性、雇佣性中解放出来，手工、脑工劳动面前人人平等，体力、智力自由发挥人人所求，每个人的手工、脑工劳动将从必然王国中解放出来转移到自由王国，每个人的体力、智力将得到全面、自由发展，每个人的生命意义将在平等与自由、人道主义与自然主义的高度统一中得到全面实现。面对脑工终结时代的来临，重构马克思机器生产工艺学批判，有助于推动人工智能科学发展与合理应用。

## 892, 通用人工智能需要在私人语言的层面上进行知识表征吗 ——来自大森庄藏的启发

徐英瑾 2020 年 11 月

复旦大学哲学学院

通用人工智能语境中的私人语言，指的是这样一个意思：表征 A 在系统甲那里的知识表征方式与同一个表征在系统乙那里的表征方式必然会有所差异。因此，在预设推论主义语义学自身有效性的前提下，A 在甲中的意义集，总会有一个子集（无论这一子集有多小）仅仅为甲自身所拥有，而无法被任何一个别的系统所拥有。因此，任何一个与甲不同的别的系统，都无法彻底地理解甲对于 A 的意义把握方式。很显然，这样一种将机器表征与哲学史上的“第一人称哲学”传统相结合的思路，是无法见容于后期维特根斯坦对于私人语言的著名反驳的。而为了与后期维特根斯坦论战，日本哲学家大森庄藏的思想资源便具有了很高的引用价值，因为他本身的哲学就可以被视为“维特根斯坦的话语方式与胡塞尔的思想内核”的日本式混合体。在对大森的哲学进行面向机器表征问题的重建的过程中，对于区域性原则与历史性原则的引入也是题中应有之义，以便为通用人工智能语境中建设私人语言的必要性提供辩护。而非公理性推理系统（纳思系统）所提供的技术手段，则会为这种想法的技术落地提供可能。

## 893, 基于卷积 LSTM 的脑机接口系统的数据增强

论文作者：Kahoko Takahashi, 2022 年 4 月 8 日

横滨市立大学、日本理化学研究所、剑桥大学、南开大学、阿里巴巴量子实验室、三峡大学经管学院等

内容：脑电图是一种非侵入性的一种检测人脑时空电信号的方法，在最近开发的脑机接口中得到了积极应用。虽然 EEG 的模式受任务影响，但其他可变因素也会影响受试者对任务的关注，并导致 EEG 信号嘈杂难以破译。为了克服这些限制，我们使用了基于人工神经网络（ANN）的方法。它们本质上对噪声具有鲁棒性，并且不需要模型。但是，它们从示例中学习并且需要大量的训练数据集。这会增加成本，需要研究的积累。为了减少网络训练所需的实验数量，我们设计了一种方法来从有限数量的训练数据集中提供人造数据。人造数据是通过在 EEG 帧上应用经验模态分解（EMD）并混合它们的本征模态函数（IMF）来生成的。我们进行了运动想象（MI）测试，参与者被要求在脑电图记录下想象左（或右）臂的运动。EEG 数据首先使用 Morlet 小波进行转换，然后由具有长短期记忆块（LSTM-RNN）的卷积神经网络处理。与标准算法相比，人造数据的引入提高了性能。即使可用的真实帧数仅为 7 或 8 个，人造数据也能提高性能。在对两个受试者（每个受试者 200 条记录）的测试中，我们对两个受试者都达到了高于 88% 的准确度。对于表现不佳的受试者，其脑电图的准确性较低，人造数据带来的改善尤其明显。想象识别准确率约为 89%，训练集有 360 帧，其中 300 帧是从 60 个真实帧开始人工创建的。我们相信这种合成人工数据的方法可能有助于开发新的、更有效的方法来训练脑机接口的神经网络。

#### **894, 自动驾驶运动中的多摄像机多重三维目标跟踪**

PhaNgyuen 等, 2022. 4. 21

美国阿肯色大学, 科罗拉多大学

目标检测和跟踪已经成为自动驾驶中最重要的任务之一。在最近的公共数据集中拥有大量数据有助于改进基于深度学习的 3D 对象检测。然而，它在实践中也带来了更具挑战性的问题，例如在各种视角和环境保持高精度和延迟性能。此外，多目标跟踪（MOT）通常与 3D 目标检测结合使用，以跟踪目标并保持跨视频帧预测的稳定性。为了处理多个视图，多摄像机多目标跟踪（MC-MOT）是一种常见方法，但同样存在一些弊端。

为了解决上述问题和挑战，本文提出一种新的 MC-MOT 框架，该框架由包含被跟踪对象外观和运动特征的节点，以及被跟踪对象或节点之间的加权边构成全局图。基于两个被跟踪对象或节点之间的外观和位置相似性计算边缘权重。其次，本文提出了一种新的自回归图变换网络，包括一个自注意力层来变换外观特征和交叉注意力机制来预测物体的运动特征。该网络有助于获得更健壮的节点嵌入，以在对象位于摄像机侧视图上时保持精确跟踪。然后，进一步使用运动传播和节点合并模块对预测结果进行后处理。最后，将使用综合评估标准对框架进行评估，以证明其与之前的 MC-MOT 框架相比的鲁棒性。在 nuScenes 数据集上的实验结果，证明了该方法在现有基于视觉的跟踪数据集上具有 SOTA 性能的优势。

## 895, 基于强化学习的产生式对抗网络的 IPv6 多模式目标生成

TianYuCui 等, 2022. 4. 22

中科院信息工程研究所, 中科大网络安全学院

由于网络速度和计算能力有限，全球 IPv6 扫描一直是研究人员面临的挑战。最近有人提出了目标生成算法，通过预测要扫描的候选集来克服互联网评估的问题。然而，IPv6 自定义地址配置出现了不同的寻址模式，阻碍了算法推断。广泛使用

的 IPv6 别名还可能误导算法，使其发现别名区域，而不是有效的主机目标。本文介绍了一种基于生成对抗网和强化学习的多模式目标生成新体系结构 6GAN。6GAN 强制多个生成器使用多类鉴别器和别名检测器进行训练，以生成具有不同寻址模式类型的非别名活动目标。来自鉴别器和别名检测器的奖励有助于监督地址序列决策过程。经过对抗性训练后，6GAN 的生成器对每个模式都能保持很强的模仿能力，6GAN 的鉴别器获得了优秀的模式识别能力，准确率为 0.966。实验表明，我们的工作通过达到更高质量的候选集而优于最先进的目标生成算法。

## 896. 一种可解释的概率自回归神经网络时间序列预测模型

MadhunimaPanja 等，2022. 4. 21

印度班加罗尔国际信息技术研究所

随着该领域要求机器学习 (ML) 模型的内部决策逻辑更加透明，目前处理可解释人工智能 (XAI) 的信息系统 (IS) 研究的数量呈爆炸式增长。然而，XAI 中包含的大多数技术都提供了事后分析解释，必须谨慎考虑，因为它们仅使用基础 ML 模型的近似值。因此，作者们的论文研究了一系列本质上可解释的 ML 模型，并讨论了它们对 IS 社区的适用性。更具体地说，作者们的重点是广义加法模型 (GAM) 的高级扩展，其中预测变量以非线性方式独立建模，以生成可以捕获任意模式但仍然完全可解释的形状函数。在作者们的研究中，作者们评估了五个 GAM 与六个传统 ML 模型相比的预测质量，并评估了它们的视觉输出以实现模型的可解释性。在此基础上，作者们调查了它们的优点和局限性，并得出了进一步改进的设计含义。

## 897, 解释黑盒模型的方法综述

RiccardoGuidotti 等, 2022. 4. 22

意大利比萨大学

近年来,许多准确的决策支持系统被构建为黑匣子,即向用户隐藏其内部逻辑的系统。这种缺乏解释既是一个实际问题,也是一个伦理问题。文献报道了许多旨在克服这一关键弱点的方法,有时以牺牲可解释性的准确性为代价。可以使用黑盒决策系统的应用程序是多种多样的,每种方法通常都是为了为特定问题提供解决方案,因此,它明确或隐含地描述了自己对可解释性和解释的定义。本文的目的是对文献中关于解释概念和黑盒系统类型的主要问题进行分类。给定一个问题定义、一个黑盒类型和一个期望的解释,这个调查应该帮助研究人员找到对他自己的工作更有用的建议。提出的开放黑盒模型方法分类也应该有助于正确看待许多研究的开放问题。

## 898, NeuroSense 神经拟态计算芯片

以色列 POLYN 科技公司, 2022. 4. 13

POLYN Technology 今天宣布,其首款神经拟态模拟信号处理器 (NASP) 芯片已完成封装和评估,证明了该技术的大脑模拟架构。这是第一个在传感器旁边使用的 Tiny AI 真正的模拟设计。POLYN Technology 是超低功耗性能 NASP 技术的创新提供商,也是独特的 Tiny AI 芯片及其相关 IP 的生产商。POLYN Technology 的首席执行官兼创始人 Aleksandr Timofeev 表示:“这一成就验证了我们跨国团队的密集工作。“我们的芯片代表了连接模拟计算和数字核心的最先进技术。它的设计考虑到了神经科学,在中心学习之前复制了人脑主要皮层

区域在外围进行的预处理”。NASP 芯片实现了传感器节点和云端之间的全数据处理分解；它真正体现了 Tiny AI 的概念。NASP 测试芯片包含几个神经网络。该芯片采用 55nm CMOS 技术实现。其设计证明了 NASP “神经元”模型以及 POLYN 开发的芯片设计自动化工具的技术可扩展性和效率。“我们的第一个芯片是通过 NASP 编译器和综合工具从经过训练的神经网络创建的，这些工具从软件数学模型模拟中生成 Netlist 和硅工程文件。我们将继续改进我们的技术以创造新一代芯片”，POLYN 首席运营官 Yaakov Milstain 说。

POLYN 预计该芯片将于 2023 年第一季度作为其首款可穿戴产品提供给客户，该芯片融合了 PPG 和 IMU 传感器，可实现最准确的心率测量以及人类活动的识别和跟踪。

**899**，将具有多重性、适应性和可塑性的峰值活动映射到生物合理的峰值神经网络中

ChengtingYu, 2022. 4. 21

浙江大学计算科学与技术学院

尖峰神经网络（SNN）被认为在生物学上更现实和节能，因为它模仿了人脑的基本机制。最近，利用深度学习框架的基于反向传播（BP）的 SNN 学习算法取得了良好的性能。然而，在这些基于 BP 的算法中，生物可解释性被部分忽略了。对于生物合理的基于 BP 的 SNN，作者在建模峰值活动时考虑了三种特性：多重性，适应性和可塑性（MAP）。在多重性方面，作者提出了一种具有多尖峰传输的多尖峰模式（MSP），以增强模型在离散时间迭代中的鲁棒性。为了实现适应性，作者在 MSP 下采用尖峰频率自适应（SFA）来减少尖峰活动，从而提高效率。对于可

塑性，作者提出了一种可训练的卷积突触，该突触模拟尖峰响应电流，以增强峰值神经元的多样性，以进行时间特征提取。所提出的 SNN 模型在神经形态数据集 N-MNIST 和 SHD 上实现了竞争性能。实验结果表明，上述 3 个方面对尖峰活动的迭代鲁棒性、尖峰效率和时间特征提取能力具有重要意义。综上所述，这项工作提出了一种可行的 MAP 生物启发尖峰活动方案，为将生物学特征嵌入尖峰神经网络提供了一种新的神经形态视角。

## 900, 使用自动挖掘知识和深度神经网络回答中医问题：端到端解决方案

Lizhang 等, 2022. 4. 15

浙江大学先进技术研究院、浙江大学计算机科学系

互联网上的医疗信息迅速增加，已成为搜索引擎使用的主要目标之一。然而，互联网上的医疗信息存在质量和可访问性问题，普通用户无法方便地获得医疗问题的答案。作为一种解决方案，研究人员构建了医学问答 (QA) 系统。然而，用中文进行医学 QA 的研究落后于基于英语的系统的工作。这种滞后主要是由于构建高质量知识库的困难和中文医学语料库的利用不足。

### 结果

本研究开发了一种端到端的解决方案，以低成本和时间实施中文医疗质量保证系统。首先，我们以几乎自动的方式从医院数据（电子健康/医疗记录）中创建了一个高质量的医学知识图谱，该图基于使用自举技术标记的数据训练了一个监督模型。然后，我们设计了一个基于记忆神经网络和注意力机制的 QA 系统。最后，我们训练系统从知识库和互联网上的 QA 语料库中生成答案。

### 结论

自举和深度神经网络技术可以从电子健康/医疗记录中构建具有令人满意的精度和覆盖率的知识图谱。我们提出的上下文桥机制执行具有各种语言特征的训练。我们的 QA 系统可以在回答主题受限的医学问题方面达到最先进的质量。正如我们评估的那样，复杂的中文语言处理技术，如分割和解析，对于实践来说不是必需的，复杂的架构也不是构建 QA 系统所必需的。最后，我们使用我们的方法创建了一个应用程序用于互联网 QA 使用。

## 901, 可池化分层图表示学习

ZhitaoYing 等, 2022. 4. 22

美国斯坦福大学、加拿大麦吉尔大学、美国南加州大学

图神经网络 (GNN) 通过有效学习节点嵌入, 彻底改变了图表示学习领域, 并在节点分类和链接预测等任务中取得了最先进的成果。然而, 当前的 GNN 方法本质上是扁平的, 并且不学习图的层次表示——这一限制对于图分类任务尤其成问题, 其目标是预测与整个图关联的标签。本文提出了 DiffPool, 这是一个可微的图池化模块, 它可以生成图的层次表示, 并且可以以端到端的方式与各种图神经网络架构相结合。DiffPool 为深度 GNN 的每一层的节点学习可微分软集群分配, 将节点映射到一组集群, 然后形成下一个 GNN 层的粗化输入。我们的实验结果表明, 与所有现有的池化方法相比, 将现有的 GNN 方法与 DiffPool 相结合, 在图分类基准上的准确率平均提高了 5-10%, 在五分之四的基准数据集上实现了新的最新技术。

## 902, 基于历史角度分析的强人工智能论争 (通用人工智能)

王彦雨, 2018. 12

中科院自然科学史研究所

在人工智能发展史上,强人工智能(“强 AI”)一直是一个争议不断但却又不断引发人们关注的议题. 对于“强 AI”理念,我们应合理看待其所发的各种争论:(1)“强 AI”理念是推动人工智能界不断打破人机界限、使 AI 技术向前发展的重要信念;(2)“强 AI”争论背后所反映的是不同社会要素,特别是“两种文化”(科学文化和人文文化)之间的张力,且 AI 界的这几种文化也经历着由对立、冲突,到逐渐的尝试性对话与合作这一过程.(3)AI 界、哲学界在对待“强 AI”这一议题的态度并非一成不变,AI 界经历了一个由乐观与支持到悲观与放弃以及现在的谨慎心态,而哲学界对于“强 AI”的态度则是沿着由批判与质疑到现在的大力宣扬这一路径演变,且当前他们对强 AI 所可能引发的风险更为忧虑.(4)“强 AI”概念需要进行重新界定,使其成为科学而非单纯的“科幻式”概念,并给予强 AI 风险议题更多关注。

### 903, 强人工智能争论过程中的“态度转换”现象研究

王彦雨, 2020. 8

中科院自然科学史研究所

在分析“强人工智能(强 AI)”概念的源起、类型及特征的基础上,对强 AI 争论过程中的“态度反转”现象进行了详细描述,即强 AI 争论过程中的两个主要参与主体——人工智能界及哲学社会科学界对于“强 AI 是否会实现”这一议题,均经历了态度上的反转(由赞同到反对抑或相反),并对“态度转换”现象背后的动力机制进行了解析. 文章最后强调,传统的强 AI 概念过于科幻,并提出基于“自主性”理念的“

强 AI2",以期破解传统强 AI 争论过程中的"不落地"困境。

#### 904, 冠状病毒病的脉电图发现

Pilato Madison 等, 2022.2

匹兹堡大学综合癫痫中心、匹兹堡大学医学中心

内容:

目的: 描述了冠状病毒病 (COVID-19) 的神经系统表现, 例如脑病和癫痫发作。据我们所知, 尚未报告 COVID-19 中的详细脑电图结果。这份报告增加了稀缺的证据。

方法: 我们确定了八名在我们医院系统中接受脑电图监测的 COVID-19 阳性患者。

结果: 脑电图最常用于意识水平的改变, 一种非特异性的神经系统表现。我们观察到所有患者的全身性背景减慢和三名患者的具有三相形态的全身性癫痫样放电。在一名有局灶性癫痫病史的患者和另一名没有此类病史的患者中观察到局灶性电图癫痫发作。8 名患者中有 5 名先前诊断为癫痫, 这表明先前存在的癫痫可能是 COVID-19 相关神经系统表现的潜在危险因素。接受脑电图检查的八名患者中有五名经历了致命的感染结果。

结论: 我们的研究结果强调了先前的观察结果, 即神经系统表现在严重病例中很常见。COVID-19 癫痫患者可能出现神经系统表现和脑电图异常的风险增加。

#### 905, 人机系统中的可解释性

AviRosenfeld 等, 2022.4.29

耶路撒冷技术学院

本文介绍了人类代理系统中可解释性的分类。我们考虑关于可解释性的原因、对象、内容、时间和方式的基本问题。首先，我们定义了可解释性，以及它与可解释性、透明度、明确性和忠实性等相关术语的关系。这些定义使我们能够回答为什么系统需要可解释性，它面向谁以及可以生成哪些解释来满足这种需求。然后我们考虑什么时候应该向用户展示这些信息。最后，我们考虑如何使用客观和主观措施来评估整个系统。最后一个问题是最全面的，因为它需要评估有关可解释性的所有其他问题。

## 906. 将糖尿病足溃疡临床描述转化为机器可解释的分割

ConnahKendrick 等, 2022. 4. 26

曼彻斯特城市大学

糖尿病足溃疡是一种需要密切监测和管理的严重疾病。为了训练机器学习方法自动描绘溃疡，临床工作人员必须提供基本事实注释。在本文中，作者们提出了一个新的糖尿病足溃疡数据集，即 DFUC2022，这是临床医生手动划定溃疡区域的最大分割数据集。作者们评估临床描述是否可以通过深度学习网络进行机器解释，或者是否应该使用图像处理精细轮廓。通过使用一系列流行的深度学习算法提供基准结果，作者们对 DFU 伤口描绘的局限性提出了新的见解，并报告了相关问题。本文提供了对基线模型的一些观察，以促进 DFUC2022 Challenge 与 MICCAI 2022 结合。排行榜将按 Dice 得分排名，其中基于 FCN 的最佳方法为 0.5708，DeepLabv3+ 获得最佳得分为 0.6277。本文表明，使用精细轮廓作为地面实况的图像处理可以提供与机器预测结果更好的一致性。

## 907, 用于神经拟态计算的可重构卤化物钙钛矿纳米晶体忆阻器

Robit Abraham John 等, 2022. 4. 19

瑞士苏黎世联邦理工学院、材料科学与技术研究所、苏黎世大学, 英国海德公司  
许多内存计算框架要求电子设备具有特定的开关特性, 以实现所需的计算复杂性水平。现有的忆阻器件无法重新配置以满足各种易失性和非易失性开关要求, 因此依赖于针对目标应用的定制材料设计, 限制了它们的通用性。结合离子扩散和漂移机制的“可重构忆阻器”可以解决这些限制, 但它们仍然难以捉摸。在这里, 我们提出了一种可重构的卤化物钙钛矿纳米晶忆阻器, 它通过可控的电化学反应实现了扩散/挥发性和漂移/非挥发性模式之间的按需切换。6 个周期) 和非易失性 ( $5.6 \times 10^3$  个周期)。我们基于对 25 个动态配置设备的 19,900 次测量, 证明了这种概念验证钙钛矿设备在具有易失性循环和非易失性读出层的基准储层网络上的相关性。

## 908, 高效、准确地转换具有突发尖峰的尖峰神经网络

YangLi 等, 2022. 4. 28

中科院自动化所脑启发智能研究中心

尖峰神经网络 (SNN) 作为一种受大脑启发的节能神经网络, 引起了研究人员的兴趣。虽然尖峰神经网络的训练仍然是一个悬而未决的问题。一种有效的方法是将训练有素的 ANN 的权重映射到 SNN, 以实现高推理能力。然而, 转换后的尖峰神经网络经常遭受性能下降和相当大的时间延迟。为了加快推理过程并获得更高的准确性, 理论上作者从 IF 和 ReLU 之间的差异, 时间维度和池化操作三个角度分析了转换过程中的误差。作者提出了一种用于释放突发峰值的神经元模型, 这

是一种廉价但高效的求解残差信息的方法。此外，提出横向抑制池（LIPooling）来解决转换过程中 MaxPooling 引起的不准确性问题。CIFAR 和 ImageNet 上的实验结果表明，作者的算法是高效和准确的。例如，作者的方法可以确保 SNN 的近乎无损的转换，并且在 0.693 下仅使用大约 1/10（小于 100）的仿真时间×能耗的典型方法。

### 909, 知识结构感知任务导向对话生成

MdRashadAIHasanRony 等, 2022. 4. 23

波恩大学、汉堡大学

面向任务的对话生成具有挑战性，因为基础知识通常是动态的，并且很难将知识有效地整合到学习过程中。在这种情况下，要同时产生类人的反应和信息丰富的反应尤其具有挑战性。最近的研究主要集中在各种知识蒸馏方法，其中知识库中事实之间的潜在关系没有被有效地捕获。在本文中，我们更进一步，展示了知识图谱的结构信息如何提高系统的推理能力。具体来说，我们提出了 DialoKG，这是一种新颖的面向任务的对话系统，可以有效地将知识整合到语言模型中。我们提出的系统将关系知识视为知识图谱，并引入（1）结构感知知识嵌入技术，以及（2）知识图谱加权注意力掩蔽策略，以促进系统在对话生成期间选择相关信息。实证评估证明了 DialoKG 在几个标准基准数据集上优于最先进方法的有效性。

### 910, 通用人工智能的数学基础初探

吕陈君、邹晓辉, 2020. 3

中美塞尔研究中心

通用人工智能 (AGI) 的数学基础，是计算机科学和神经科学这两种途径之外的第三种研究途径或方法，其核心是区分计算机的算术范式与学习机的统计范式乃至理解机的几何范式以及贯穿三者的逻辑范式，三类机器之间实现不同数学范式的相互映射和相互转换。其目的是探寻 AGI 的数学基础，这是智能机自动编程和理解各类模式的关键。如果通用学习机是 AGI 的基石，那么通用理解机则是理解毕达哥拉斯、康德和哥德尔等深刻哲思的关键。其意义在于：明确了探讨三类机器的统一数学原理的问题与探寻 AGI 的数学基础之间的关系，同时明确和验证了人工智能和计算机科学之间的关系。

### 911, 长期自然状态人类颅内神经活动和姿势记录

Steven M. Peterson 等, 2022. 4. 21

华盛顿大学

了解自然场景中人类运动的神经基础对于将神经科学研究扩展到受限的实验室范式之外至关重要。在这里，我们为 12 名人类参与者建立了在长期皮质脑电图的注释关节 (AJILE12) 数据集。它是最大的公开可用的人类神经行为数据集。该数据集是在临床癫痫监测期间随机记录的。AJILE12 包括跨 55 个半连续天的自然运动的同步颅内神经记录和上身姿势轨迹，以及相关元数据，包括数千个手腕运动事件和带注释的行为状态。在总共 1280 小时的时间里，每个参与者至少有 64 个电极以 500 Hz 的频率提供神经记录，并从 1.18 亿个视频帧估计了 9 个上身关键点的姿势轨迹。为了促进数据探索和重用，我们在神经数据无国界 (NWB) 数据标准中的 DANDI 存档上共享了 AJILE12。

## 912, 预测机器剩余使用寿命的可解释回归框架

Talhat Khan 等, 2022. 4. 28

Quaid-e-Awam 工程技术科学大学, 蒙斯特科技大学、亚洲大学

预测机器的剩余使用寿命 (RUL) 是预测性维修的关键任务之一。该任务被视为一个回归问题, 机器学习 (ML) 算法用于预测机器部件的 RUL。这些 ML 算法通常被用作一个黑匣子, 完全关注性能, 而不确定算法决策背后的潜在原因及其工作机制。我们认为, 单凭绩效 (就均方误差 (MSE) 等而言) 不足以建立利益相关者对 ML 预测的信任, 而是需要更多关于预测背后原因的见解。为此, 在本文中, 我们通过提出一个用于预测机器 RUL 的可解释回归框架来探索可解释人工智能 (XAI) 技术的潜力。我们还评估了几种 ML 算法, 包括基于经典和神经网络 (NNs) 的任务解决方案。对于解释, 我们依赖两种模型不可知的 XAI 方法, 即局部可解释模型不可知解释 (LIME) 和 Shapley 加法解释 (SHAP)。我们相信, 这项工作将为该领域的未来研究提供基线。

## 913, 基于磁场诱导 Skymion 动力学的神经形态计算模式识别

Tomoyuki Yokouchi 等, 2022. 4. 27

里肯紧急物质科学中心, 东京大学, 亚当-米维兹大学

物理系统中的非线性现象可以用于低能耗的大脑启发计算。拓扑自旋结构 skyrmion 的动力学响应是这种神经形态计算的候选之一。然而, 它的能力还没有得到很好的实验探索。在这里, 我们用磁场诱导的 Skymion 动力学产生的非线性响应来实验演示神经形态计算。我们设计了一个结构简单的基于 skyrmion 的

神经形态识别装置，并成功地实现了手写数字识别，识别准确率高达 94.7%，波形识别也取得了成功。值得注意的是，识别准确度与设备中 Skyrmion 的数量之间存在正相关。skyrmion 系统的大自由度，如位置和大小，产生了更复杂的非线性映射和更大的输出维数，因此具有更高的精度。我们的研究结果为开发节能和高性能的 skyrmion 神经形态计算设备提供了指导。

#### 914, 脑机接口：研究从视觉诱发电位到纯粹想象稳态电位的转变

Arturo Micheli 等，

都灵理工大学、大西洋高等国立矿业电信学校

基于稳态视觉诱发电位（SSVEP）的脑-机接口（BCI）已被证明是有效的，并提供了显著的准确性和信息传输率。然而，这一系列策略需要外部设备来提供该技术所需的频率刺激。这限制了它们的应用场景，尤其是与其他 BCI 方法相比。在这项工作中，我们研究了基于 SSVEP 诱发刺激的纯视觉想象在 EEG 输出中获得频率响应的可能性。我们的研究表明，不仅 EEG 信号呈现出与用户关注的频率相关的特定频率峰值，而且还可以实现令人满意的分类精度，为稳健可靠的视觉图像 BCI 模式铺平道路。

#### 915, 联网车辆的端到端协同感知驾驶

Jiaxon Cui 等，2022.5.5

得克萨斯大学奥斯汀分校，斯坦福大学，索尼 AI 部门

在过去几年中，用于自动驾驶车辆的光学传感器和学习算法有了显著的进步。尽管如此，由于视线感应能力有限，以及数据驱动方法在处理极端情况时的脆弱性，

自动驾驶汽车的可靠性受到了阻碍。随着近年来通信技术的发展，车对车通信的协同感知已成为在危险或紧急情况下增强自动驾驶能力的一种很有前景的范例。本文提出了 COOPERNAUT，这是一种端到端的学习模型，它将跨车辆感知用于基于视觉的合作驾驶。该模型将激光雷达信息编码为紧凑的基于点的表示，可以通过真实的无线信道在车辆之间作为消息传输。COOPERNAUT 会学习在真实的 V2V 信道容量下附近车辆共享的编码激光雷达信息。为了在符合带宽限制的情况下传达附近车辆的有意义的场景信息，作者设计了基于 Point Transformer 的驾驶策略架构，这是一个用于点云处理的自注意力网络。该架构将每个联网车辆上的原始点云预处理为空间感知神经表示。这些表示是紧凑的，可以通过真实的无线信道高效地传输。同时，它们在物理上是固定的，因此可以通过自我表征进行空间转换和聚合。

为了评估模型，作者还开发了一个基于 CARLA 的模拟框架 AUTOCASTSIM，这是一个网络增强的驾驶模拟框架，设计了三种容易发生事故的场景，所有场景的设计都是为了挑战自我认知，让模型完全理解交通状况。AUTOCASTSIM 有一个内置的网络模拟，用于定制多车辆通信，以及一个具有特权信息的专家驾驶模型，并使用基于体素的基线和不同的传感器融合方案来评估 COOPERNAUT。在 AUTOCASTSIM 上的实验表明，在具有挑战性的驾驶环境中，本文提出的合作感知驾驶模型比以自我为中心的驾驶模型的平均成功率提高了 40%，并且比之前提出的 V2VNet 模型的工作带宽需求减少了 5 倍。

## 916. 基于强化学习的快速移动

GabrielBMargolis 等，2022.5.6

麻省理工学院

对于腿型机器人来说，在野外进行短跑和高速转弯等敏捷动作是一项挑战。我们提供了一个端到端学习控制器，实现了麻省理工学院迷你猎豹创纪录的敏捷性，维持速度高达 3.9 米/秒。该系统在草地、冰和砾石等自然地形上快速运行和转向，并对干扰做出了稳健的响应。我们的控制器是通过强化学习在仿真中训练的神经网络，并转移到现实世界中。这两个关键组成部分是 (i) 关于速度指令的自适应课程和 (ii) 利用之前的工作为 sim-to-real 传输提供在线系统识别策略。

### 917, 窥视黑盒子：一项关于可解释人工智能 (XAI) 的调查

AminaAdadi, 2022. 5. 6

摩洛哥西迪-阿卜杜勒大学

在第四次工业革命的曙光之际，我们目睹了人工智能 (AI) 在日常生活中的快速和广泛采用，这有助于加速向更加算法化的社会转变。然而，即使有了如此前所未有的进步，使用基于人工智能的系统的一个关键障碍是它们通常缺乏透明度。事实上，这些系统的黑盒特性允许进行强有力的预测，但无法直接解释。这个问题引发了一场关于可解释人工智能 (XAI) 的新辩论。一个研究领域对提高基于人工智能的系统的信任度和透明度有着巨大的希望。它被认为是人工智能在不受干扰的情况下继续稳步发展的必要条件。该调查为感兴趣的研究人员和从业者提供了一个切入点，以了解与 XAI 相关的年轻且快速发展的研究机构的关键方面。通过文献的镜头，我们回顾了有关该主题的现有方法，讨论了围绕其领域的趋势，并提出了主要的研究轨迹。

## 918, 基于心电数据检测心律失常的可解释机器学习模型的开发

Shourya Verma, 2022. 5. 6

海德堡大学科学计算跨学科中心

因为心电图 (ECG) 信号的分析是由心脏病专家手动执行的, 非常耗时。因此, 越来越多地提出通过机器学习 (ML) 分类实现自动化, 这将允许 ML 模型学习心跳的特征并检测异常。缺乏可解释性阻碍了深度学习在医疗保健中的应用。本论文基于最先进的模型构建了卷积神经网络 (CNN) 和长短期记忆 (LSTM) 分类器, 并将它们的性能和可解释性与浅层分类器进行了比较。在这里, 利用全局和局部可解释性方法来了解整个数据集中因变量和自变量之间的相互作用, 并分别检查每个样本中的模型决策。部分依赖图、Shapley 加法解释、排列特征重要性和梯度加权类激活图 (Grad-Cam) 是在时间序列 ML 模型上实现的四种可解释性技术, 用于对 ECG 节律进行分类。特别是, 作者利用 Grad-Cam, 这是一种局部可解释性技术, 并检查其可解释性是否在每个类别中正确和错误分类的 ECG 节拍之间变化。此外, 分类器使用 K-Fold 交叉验证和 Leave Groups Out 技术进行评估, 作者使用非参数统计测试来检查是否存在重大差异。研究发现, Grad-CAM 是解释提出的 CNN 和 LSTM 模型预测的最有效的可解释性技术。作者得出结论, 所有高性能分类器在进行预测时都会查看心电图节律的 QRS 复合波。

## 919, 在 SpiNNaker 上的 iCub 仿人机器人的事件驱动仿生注意力系统

Giulia Veronica D'Angelo 等, 2022. 4. 28

英国曼彻斯特大学, 意大利技术研究院

注意力引导观察者注视有趣的项目, 只允许对场景的选定区域进行详细分析。机

机器人可以利用场景中特征的感知组织来引导其注意力，从而更好地理解其环境。当前自下而上的注意力模型与标准 RGB 摄像头配合使用，需要大量时间以基于帧的方式检测最显著的项目。事件驱动相机是一种创新技术，能够以高时间分辨率和低延迟异步检测场景中的对比度变化。我们提出了一种新的神经拟态管道，利用事件驱动相机的异步输出生成场景的显著性贴图。为了进一步减少潜伏期，神经拟态注意模型在 SpiNNaker（一个专用的神经拟态平台）上的一个尖峰神经网络中实现。已将提议的实现与其仿生 GPU 对应物进行了比较，并根据地面真相固定地图进行了基准测试。该系统成功地检测到场景中的项目，生成与 GPU 实现相当的显著性贴图。异步管道的平均延迟为 16 毫秒，以生成可用的显著性图。

## 920, 尖峰卷积网络

ZulunZhu 等, 2022. 5. 5

新加坡南洋理工大学

图卷积网络（GCN）由于在学习图信息方面具有出色的表示能力，因此实现了令人印象深刻的性能。然而，当在深度网络上实现 GCN 时，需要昂贵的计算能力，这使得它们难以部署在电池供电的设备上。相比之下，执行生物保真推理过程的峰值神经网络（SNN）提供了一种节能的神经架构。在这项工作中，作者提出了 SpikingGCN，这是一个端到端框架，旨在将 GCN 的嵌入与 SNN 的生物保真特性相结合。原始图形数据基于图卷积的合并被编码为尖峰序列。作者通过利用与神经元节点相结合的完全连接的层来进一步模拟生物信息处理。在广泛的场景中（例如引文网络，图像图分类和推荐系统），作者的实验结果表明，与最先进的方法相比，所提出的方法可以获得有竞争力的性能。此外，作者表明，神经形态芯片

上的 SpikingGCN 可以为图数据分析带来明显的能效优势，这证明了其构建环境友好型机器学习模型的巨大潜力。

## 921, 关于图像分类的端到端的深度学习架构

MuhanZhang 等, 2022. 3. 30

华盛顿大学

本文提出了一种接受任意结构图的新型神经网络结构，主要解决了图分类问题中的两个挑战：

- (1) 如何提取有用的特征来描述那些编码在图中的丰富信息，以达到分类的目的；
- (2) 如何以有意义且一致的顺序读取图。

从实际的操作来看，这两个 challenge 分别对应图卷积以及图池化。对此，为了解决第一个挑战，文中设计了一个局部图卷积模型，并给出了它与两个图核的关系；为了解决第二个挑战，设计了一个新的 SortPooling 层，以一致的顺序对图顶点进行排序，这样传统的神经网络就可以在图上进行训练；并通过 top-k 选取对卷积之后的图进行池化。

本文的贡献如下：

- (1) 提出了一种新颖的端到端深度学习结构用于图分类。它直接接受图作为输入，不需要任何预处理。
- (2) 提出了一种新的空域图卷积层来提取多尺度的顶点特征，并与目前流行的图核进行了类比来解释其工作原理。
- (3) 提出了一个新的 SortPooling 层来对顶点特征进行排序，而不是对它们进

行汇总,这样可以保留更多的信息,并允许我们更够从全局图的拓扑结构中学习。

(4) 在基准图分类数据集上的实验结果表明,深度图卷积神经网络(DGCNN)与先进的图核相比具有很强的竞争力,并且在图分类方面明显优于许多其他深度学习方法。

## 922, 通过系统回顾定义知识图谱开发过程

GyteTamasauskaite 等, 2022. 4. 30

荷兰阿姆斯特丹大学

知识图谱在工业中被广泛使用,并在学术界进行了研究。但是,在知识图谱开发中应用的模型各不相同。分析和提供知识图谱开发常用方法的综合将使研究人员和从业者更好地理解所涉及的整体过程和方法。因此,本文旨在定义知识图谱开发的整个过程及其关键组成步骤。为此,对文献进行了系统回顾和概念分析。将所得过程与案例研究进行比较,以评估其适用性。提议的过程提出了一种统一的方法,并在构建和管理知识图谱时为研究人员和从业者提供指导。

## 923, 如何设计具有自主意图的人工智能体——一项基于安斯康意图哲学的跨学科探索

徐英瑾, 2018. 11

复旦大学哲学学院

具有自主意图、只依赖小数据运作的通用人工智能系统的出现,并不会像有些人所预估的那样导致"机器奴役人类"的局面出现,因为此类技术对于小数据的容忍可以大大增加此类技术的潜在用户的数量,并使得体现不同用户价值观的通用人

工智能系统能够大量出现。这样一来,具有不同意图的通用人工智能系统彼此之间的对冲效应,最终会使得任何一种具有特定意图的通用人工智能系统都无法占据主宰地位。相反,由于作为专用人工智能技术代表的深度学习技术的运用在原则上就需要大量数据的喂入,其对于民众隐私权的侵犯就成为一种难以被全面遏制的常态,因此,此类技术的发展在原则上就会加强一部分技术权贵对于大多数民众的统治地位。不过,要在通用人工智能系统里实现对于意图的工程学建模,就需要我们在哲学层面上首先厘清关于意图的种种哲学迷思。在这个问题上,美国女哲学家安斯康的意图理论是一个比较好的讨论起点。具体而言,安斯康关于"意图是在欲望驱使下做某事的理由"的观点,是可以在通用人工智能的语境中被实现的,但是她关于信念与意图之二元对立的观点,却在不少地方有失偏颇。而"非公理化推理系统"(纳思系统),则将为吸纳安斯康意图论的合理部分提供相应的工程学手段。

#### 924. 基于统计前向规划算法的游戏通用人工智能

LucasSimon 等, 2019.9

伦敦玛丽女王大学, 中科院自化研究所, 中科院大学, 青岛智能产业技术研究院  
统计前向规划 (statistical forward planning, SFP) 算法使用仿真模型 (也称为前向模型) 自适应地搜索有效的动作序列, 此类算法提供了一种简单通用的方法, 为各种游戏提供快速自适应的 AI 控制。介绍了两种常用的 SFP 算法: 蒙特卡罗树搜索和滚动层进化, 并证明了在没有任何事先训练的情况下, SFP 算法就可以在各种视频游戏中出色地运行。总体来说, 统计前向规划方法是一种使用前向模型来模拟可能的未来状态的方法。这些方法是一组稳健的、通用的、随机的人工

智能算法，在很多问题上都取得了很好的效果，同时它适应性很强，只需要快速地复制前向模型就可以取得很好的泛化性能。采用错误的模型参数进行的实验也说明，不是完美的模型才有效。

## 925, 脑机接口

迈向被动 BCI: 模拟驾驶舱场景中对方问和角度特定扰动的神经响应的异步解码  
ShayanJalilpour 等, 2022. 4. 26

奥地利格拉茨科技大学神经工程研究所

神经影像学研究提供了证据，证明失去平衡会在脑电图 (EEG) 中引起特定的神经瞬态波复合物，称为扰动诱发电位 (PEP)。在线解码来自正在进行的 EEG 信号的平衡扰动可以建立实现被动脑机接口 (pBCI) 作为航空/驾驶辅助系统的一部分的可能性。在这项研究中，我们研究了使用 EEG 信号识别四种不同条件下扰动的存在和表达的可行性。15 名健康参与者坐在滑翔机驾驶舱中时经历了四种不同的姿势变化。连接到滑翔机的机器人暴露突然的扰动，并以 5 度和 10 度的倾斜角向左右方向移动。扰动发生在一个奇怪的范式中，参与者不知道扰动的时间和表达。我们采用分层方法来分离扰动和静止，然后区分扰动的表达。BCI 系统的性能通过使用分类准确度和 F1 分数来评估。异步地，我们实现了 89.83% 和 73.64% 的平均准确率，以及二进制和多类分类的平均 F1 分数分别为 0.93 和 0.60。这些结果证明了 pBCI 在现实情况下检测平衡扰动的实用性。BCI 系统的性能通过使用分类准确度和 F1 分数来评估。异步地，我们实现了 89.83% 和 73.64% 的平均准确率，以及二进制和多类分类的平均 F1 分数分别为 0.93 和 0.60。这些结果证明了 pBCI 在现实情况下检测平衡扰动的实用性。BCI 系统的

性能通过使用分类准确度和 F1 分数来评估。异步地，我们实现了 89.83% 和 73.64% 的平均准确率，以及二进制和多类分类的平均 F1 分数分别为 0.93 和 0.60。这些结果证明了 pBCI 在现实情况下检测平衡扰动的实用性。

## 926, 通过非侵入性脑机超平台实现人脑直接无线通信

Qian Ma 等, 2022.4.30

东南大学, 华南理工大学, 广州琶洲实验室, 新加坡国立大学, 北京大学信息科技学院, 英国格拉斯哥大学

受基于 BCI 的神经系统损伤和截肢康复技术的启发, 作者提出了一种电磁脑 - 计算机 - 超表面 (EBCM) 范式, 由人类的认知通过脑信号直接和非侵入性地进行调节。作者通过实验表明, 他们的 EBCM 平台可以将人类的思想从基于 P300 的脑电图的诱发电位无创地转化为电磁域中的数字编码信息, 这些信息可以通过信息超表面以自动化和无线方式进一步处理和传输。两个 EBCM 操作员之间通过文本传输准确的进行人脑的直接无线通信。此外, 使用相同的 EBCM 平台展示了其他几个概念验证的精神控制方案, 如视觉光束扫描、波调制和模式编码, 展示了灵活定制的信息处理和合成能力。