

# 评人工智能如何走向新阶段 (兼谈国内外AI跟贴评论)

陆首群

2023.1.5

国内外AI跟贴留言  
(1001条~1145条)

第十四集

中国开源软件推进联盟  
China OSS Promotion Union

# 评人工智能如何走向新阶段？

陆首群

2023. 1. 5

我们一直关注：基于大规模语义网络（知识图谱）以实现认知智能这条走向强人工智能的路径。我们在第十一集的点评中指出：采用数据、知识双驱动，立足于知识工程，研发大规模语义网络（知识图谱）以支持实现认知智能，这条通向强人工智能之路难度较大，主要原因在于目前研发的语义网络缺乏逻辑推理机制，机器尚难识别常识等短板。

本期（十四集 1031 条）发表浙江大学网络空间安全研究所、浙江工业大学信息工程学院的一篇文章：《基于动态关系置信度的知识图谱重要路径发现》，文中谈到：现有的知识图谱大多是不完整的，可以通过一些推理算法进行补充。基于路径特征的推理方法由于具有较强的可解释性，在知识推理和补全领域得到了广泛的应用。但是，基于路径特征的推理方法仍存在以下几个方面的问题：路径搜索效率低，稀疏任务的路径不足，一些路径对推理任务没有帮助。

针对上述问题，本文提出一种 PC-Patch 方法，结合动态关系置信度等指标对路径特征进行评价，进而指导路径搜索，最后进行关系推理。

点评者认为，对此有兴趣者可深入了解该文作者在处理推理算法方面的论文及其实验效果。

本文集对谷歌、DeepMind 提出的 AlphaFold (涉及基因医疗、生命科学的对蛋白质结构预测和训练) 一直进行追踪报导, 在本期(十四集) 中又连续报导 1050 条 (FastFold)、1041 条、1057 条、1069 条、1075 条、1097 条、1114 条 (共 7 条), 今天 AlphaFold 的研发单位扩大了, 除谷歌、DeepMind 外, 华盛顿大学、DeepMind, London, UK, 欧洲 MolecularLaboratory, 欧洲 Bioinformatics 研究所, Hinxton, UK 等。我曾建议国内有关单位可与他们进行合作。在上述 7 条中的 1040 条, DeepMind 团队研发了 FastFold, 将 AlphaFold 训练蛋白质时间从 11 天减少到 67 小时(减少训练和推理内耗提高准确预测蛋白质结构效率)。

关于打破机器学习黑盒子实现可解释性人工智能方面, 我们是全球最早追踪报导的并经常与国内外研发单位进行深入讨论的。可解释性问题近几年一直是全球研发的亮点, 但对可解释性案例如何评估其算法或解决方案的有效性、准确性、可靠性、可信性也一直有争论, 本期在关于如何建立可解释性模型和案例分析方面, 推荐大家研读下列论文:

① 1016 条, 哈佛大学和卡耐基梅隆大学共同撰写的文章 OpenXAI: 实现模型解释的透明评估;

② 1006 条, 杜伊斯堡-埃森大学 (University of Duisburg-Essen) 的论文: 在自然对话中解释机器学习模型;

③ 1085 条，佐治亚理工学院、华盛顿大学的论文：可解释 NLP 数据集的回顾确定了 65 个具有三类主要文本解释的数据集（亮点、自由文本和结构化）。

## 国内外 AI 跟帖留言 (1001-1145)

1001, 用于 EEG 脑机接口的带有地形表示模块的卷积神经网络

Xinbin Liang, Yaru Liu, Yang Yu, Kaixuan Liu, Yadong Liu, Zongtan Zhou,  
2022. 8. 23

国防科技大学

卷积神经网络 (CNN) 在脑机接口 (BCI) 领域显示出巨大的潜力, 因为它们能够直接处理原始脑电图 (EEG) 而无需人工特征提取。原始脑电信号通常展现为由通道和时间点组成的二维 (2-D) 矩阵, 却忽略了脑电的空间拓扑信息。本文的目标是使以原始脑电信号为输入的 CNN 具有学习脑电空间拓扑特征的能力, 并在基本保持原始结构的同时提高其分类性能。所用到的方法是提出了一个脑电图拓扑表示模块 (TRM), 该模块包括: (1) 从原始 EEG 信号到 3-D 地形图的映射块; (2) 从地形图到与输入大小相同的输出的卷积块。本文将 TRM 嵌入到 3 个广泛使用的 CNN 中, 并在 2 种不同类型的公开可用数据集上对其进行了测试。结果表明, 使用 TRM 后, 3 个 CNN 的分类精度在 2 种数据集上都有所提高。DeepConvNet、EEGNet 和 ShallowConvNet 在模拟驾驶数据集 (EBDSDD) 上的平均分类准确率分别提高了 4.70%、1.29% 和 0.91%。这样做的意义在于通过使用 TRM 挖掘 EEG 的空间拓扑特征, 提高了 3 个 CNN 在 2 个数据集上的分类性能。

1002, 可解释的人工智能在网络安全应用中的最新研究

Zhibo Zhang, Hussam Al Hamadi, Ernesto Damiani 等, 2022. 8. 31

哈利法大学、扎耶德大学

本文全面回顾了目前关于网络安全应用的可解释人工智能（XAI）的方法。近年来，随着互联网和人工智能的快速发展，包括机器学习（ML）和深度学习（DL）在内的人工智能广泛应用于入侵检测、恶意软件检测、垃圾邮件过滤等网络安全领域。虽然与传统的基于签名和规则的网络安全策略相比，基于人工智能的网络攻击、威胁检测、防御方法更加先进和高效，但大多数基于 ML 的技术和 DL 的技术都是以黑匣子方式部署，安全专家无法向客户解释此类程序是如何得出特定结论的。现有人工智能技术的透明度和可解释性不足等问题，可能会降低用户对防御网络攻击模型的信心。因此，当 XAI 应用于网络安全，能在保持高精确度的同时，创建更多可解释的模型，也使用户能理解、信任和管理下一代网络防御机制。

### **1003, 受人脑皮质启发，减少多核神经形态处理器中内存资源的使用**

Vanessa R. C. Leite, Zhe Su, Adrian M. Whatley, Giacomo Indiveri, 2022. 8. 29

苏黎世大学、苏黎世联邦理工学院

受大脑启发的基于事件的神经形态处理系统，已经成为一种前景极佳的技术，特别是在生物医学电路和系统方面。然而，神经网络的神经形态和生物实现都有能量和记忆限制。为了在多核神经形态处理器中，尽量减少内存资源的使用，本文提出了一种受生物神经网络启发的网络设计方法，并且利用该方法设计了一种适用于小世界网络优化的新路由方案。同时，提出了一种硬件感知的放置算法，能优化小世界网络模型的资源分配。此外，还通过使用规范的小世界网络验证了该算法，并呈现了从中衍生的其他网络的初步结果。

#### 1004, SemSegDepth: 语义分割和深度补全的组合模型

Juan Pablo Lagos, Esa Rahtu, 2022. 9. 1

坦佩雷大学 (Tampere University)

对自主机器的性能而言，场景的整体理解是很重要的。在本文中，提出了一种端到端的新模型——SemSegDepth，用于整合语义分割和深度补全。目前，绝大多数的方法都将语义分割和深度补全作为独立的任务，而该方法依赖 RGB 和稀疏深度作为模型的输入，并生成密集深度图和相应的语义分割图像。它由特征提取器、深度完成分支、语义分割分支、共同处理语义和深度信息的联合分支组成。在 Virtual KITTI 2 数据集上进行的实验，进一步证明了在多任务网络中，整合语义分割和深度完成两种任务可以有效地提高各个任务的性能。

#### 1005, 通过脑机接口检测情绪对驾驶的影响

Mario Quiles Pérez, Enrique Tomás Martínez Beltrán, Sergio López Bernal, Alberto Huertas Celdrán, Gregorio Martínez Pérez, , 2022. 8. 30

穆尔西亚大学 (University of Murcia)、苏黎世大学 (University of Zurich)

目前，已经有许多技术用于预防交通事故，其中脑机接口 (BCI) 是最有前景的技术之一。BCI 被用于检测与司机的决定存在直接联系的场景，包括情绪状态、注意力问题、压力情况等。但当前还没有太多文献去研究 BCI 用于检测驾驶场景中的受试者情绪。在本文中，有一些问题需要解决：一是执行驾驶任务对情绪检测的影响；二是哪些情绪在驾驶场景中更容易检测。同时，文中提出了一个框架，可以利用机器学习和深度学习算法的脑电图来检测情绪。此外，还设计了一个用例，提出了两个场景：第一个场景把听声音作为主要任务，第二个场景使用驾驶

模拟器作为主要任务，将听声音作为次要任务。通过这种方式，旨在证明 BCI 在该驾驶场景中是否有用。最终，改善了现有文献的结果，对两种情绪（非刺激和愤怒）检测的准确率达到 99%，对三种情绪（非刺激、愤怒和中性）的准确率达到 93%，对四种情绪（非刺激、愤怒、中性和喜悦）的准确率达到 75%。

**1006**，在自然对话中解释机器学习模型——通过建立一个对话式的 XAI 代理

Van Bach Nguyen, Jörg Schlötterer, Christin Seifert, 2022. 9. 6

杜伊斯堡-埃森大学 (University of Duisburg-Essen)

可解释人工智能 (XAI) 的目标是设计一种方法，提供对经验模型推理过程的见解，便于人类解释它们。社会科学研究指出，这种解释是对话形式的。本文中，展示了如何将 XAI 纳入对话式代理中，使用包括自然语言理解和生成组件的代理的标准设计。首先，需要建立一个 XAI 问题库，通过把控释义的质量来扩展这个问题库，以理解用户的信息需求。随后，通过系统地调查文献，寻找合适的解释方法，提供相关信息，并列出一个全面的建议清单。这将成为机器学习模型与解释代理迈向自然对话的第一步。XAI 的相关问题和解释方法可以支持研究者提供必要的信息，以此满足用户的需求。

**1007**，保护峰值：关于尖峰神经网络对对抗性示例的可转移性和安全性

Nuo Xu, Kaleel Mahmood, Haowen Fang, Ethan Rathbun, Caiwen Ding, Wujie Wen,  
2022. 9. 7

理海大学、康涅狄格大学

尖峰神经网络 (SNNs) 因其高能效和分类性能的最新进展而备受关注。然而，与

传统的深度学习方法不同，SNNs 在对抗性例子的鲁棒性上分析和研究相对不发达。本文中，通过实验和分析三个重要的 SNNs 安全属性来推进对抗性机器学习领域。首先，展现了对 SNNs 的成功白盒对抗性攻击高度依赖于底层的代理梯度技术。其次，分析了由 SNNs 和其他最先进的架构（如 Vision Transformers 和 Big Transfer CNNs）产生的对抗性例子的可转移性。文中证明了 SNNs 并没有被 Vision Transformers 和某些类型的 CNNs 产生的对抗性例子经常欺骗。最后，开发了一种新的白盒攻击，它产生的对抗性例子能够同时骗过 SNN 模型和非 SNN 模型。文中的实验和分析涵盖了两个数据集（CIFAR-10 和 CIFAR-100），五个不同的白盒攻击和十二个不同的分类器模型。

**1008, SUNet: 用于全景分段的具有规模意识的统一网络**

Weihaoyan, Yeqiang Qian, Chunxiang Wang, Ming Yang, 2022.9.7

电气电子工程师学会 (IEEE)

全景分割结合了语义分割和实例分割的优点，可以为智能车辆提供像素级和实例级的环境感知信息，但它在分割各种尺度的物体上存在问题，特别是在极度大和小的物体上。本文中，提出了两个轻量级模块来缓解该问题。首先，Pixel-relation Block 被设计用来模拟大规模事物的全局环境信息，它是基于与查询无关的一个公式，并带来小的参数增量。其次，构造对流网络来收集小规模事物的额外高分辨率信息，为下游的分割分支提供更合适的语义特征。基于这两个模块，文中提出了一个端到端的尺度感知统一网络 (SUNet)，它对多尺度物体的适应性更强。Cityscapes 和 COCO 上的许多实验也证明了该方法的有效性。

### 1009, 一种新型的可用于主体转移脑机接口的半监督元学习方法

Jingcong Li, Fei Wang, Haiyun Huang, Feifei Qi, Jiahui Pan, 2022. 9. 7

华南师范大学、广东金融学院、琶洲实验室

脑机接口 (BCI) 提供了人脑与外部设备之间的直接通信路径。在新受试者使用 BCI 之前, 通常需要校准程序。由于主体间和主体内的差异如此之大, 导致现有主体训练的模型在新主体上表现不佳。因此, 有效的主体转移和校准方法很重要。本文中, 提出了一种半监督元学习 (SSML) 方法, 用于 BCI 的主体转移学习。首先, 拟议的 SSML 利用现有科目学习元模型。然后以半监督学习方式对模型进行微调, 使用少量标记和许多未标记的样本进行校准。这对于 BCI 应用非常重要, 因为在这些应用中, 标记的数据稀缺且昂贵, 而未标记的数据很容易获得。为了验证 SSML 方法, 测试了三种不同的 BCI 范式: 事件相关电位检测; 情绪识别; 睡眠分期。在前两个范式上, SSML 取得了超过 15% 的改善, 在第三个范式上, SSML 取得了 4.9% 的改善。结果证明, SSML 方法在 BCI 应用中的很有效。

### 1010, 在自然对话中解释机器学习模型——通过建立一个对话式的 XAI 代理

Van Bach Nguyen, Jörg Schlötterer, Christin Seifert, 2022. 9. 6

杜伊斯堡-埃森大学 (University of Duisburg-Essen)

可解释人工智能 (XAI) 的目标是设计一种方法, 提供对经验模型推理过程的见解, 便于人类解释它们。社会科学研究指出, 这种解释是对话形式的。本文中, 展示了如何将 XAI 纳入对话式代理中, 使用包括自然语言理解和生成组件的代理的标准设计。首先, 需要建立一个 XAI 问题库, 通过把控释义的质量来扩展这个问题库, 以理解用户的信息需求。随后, 通过系统地调查文献, 寻找合适的解释

方法，提供相关信息，并列出一个全面的建议清单。这将成为机器学习模型与解释代理迈向自然对话的第一步。XAI 的相关问题和解释方法可以支持研究者提供必要的信息，以此满足用户的需求。

### **1011, 保护峰值：关于尖峰神经网络对对抗性示例的可转移性和安全性**

Nuo Xu, Kaleel Mahmood, Haowen Fang, Ethan Rathbun, Caiwen Ding, Wujie Wen,  
2022. 9. 7

理海大学、康涅狄格大学

尖峰神经网络（SNNs）因其高能效和分类性能的最新进展而备受关注。然而，与传统的深度学习方法不同，SNNs 在对抗性示例的鲁棒性上分析和研究相对不发达。本文中，通过实验和分析三个重要的 SNNs 安全属性来推进对抗性机器学习领域。首先，展现了对 SNNs 的成功白盒对抗性攻击高度依赖于底层的代理梯度技术。其次，分析了由 SNNs 和其他最先进的架构（如 Vision Transformers 和 Big Transfer CNNs）产生的对抗性示例的可转移性。文中证明了 SNNs 并没有被 Vision Transformers 和某些类型的 CNNs 产生的对抗性示例经常欺骗。最后，开发了一种新的白盒攻击，它产生的对抗性示例能够同时骗过 SNN 模型和非 SNN 模型。文中的实验和分析涵盖了两个数据集（CIFAR-10 和 CIFAR-100），五个不同的白盒攻击和十二个不同的分类器模型。

### **1012, SUNet：用于全景分段的具有规模意识的统一网络**

Weihao Yan, Yeqiang Qian, Chunxiang Wang, Ming Yang, 2022. 9. 7

电气电子工程师学会（IEEE）

全景分割结合了语义分割和实例分割的优点，可以为智能车辆提供像素级和实例级的环境感知信息，但它在分割各种尺度的物体上存在问题，特别是在极度大和小的物体上。本文中，提出了两个轻量级模块来缓解该问题。首先，Pixel-relation Block 被设计用来模拟大规模事物的全局环境信息，它是基于与查询无关的一个公式，并带来小的参数增量。其次，构造对流网络来收集小规模事物的额外高分辨率信息，为下游的分割分支提供更合适的语义特征。基于这两个模块，文中提出了一个端到端的尺度感知统一网络（SUNet），它对多尺度物体的适应性更强。Cityscapes 和 COCO 上的许多实验也证明了该方法的有效性。

### 1013, 一种新型的可用于主体转移脑机接口的半监督元学习方法

Jingcong Li, Fei Wang, Haiyun Huang, Feifei Qi, Jiahui Pan, 2022.9.7

华南师范大学、广东金融学院、琶洲实验室

脑机接口（BCI）提供了人脑与外部设备之间的直接通信路径。在新受试者使用 BCI 之前，通常需要校准程序。由于主体间和主体内的差异如此之大，导致现有主体训练的模型在新主体上表现不佳。因此，有效的主体转移和校准方法很重要。本文中，提出了一种半监督元学习（SSML）方法，用于 BCI 的主体转移学习。首先，拟议的 SSML 利用现有科目学习元模型。然后以半监督学习方式对模型进行微调，使用少量标记和许多未标记的样本进行校准。这对于 BCI 应用非常重要，因为在这些应用中，标记的数据稀缺且昂贵，而未标记的数据很容易获得。为了验证 SSML 方法，测试了三种不同的 BCI 范式：事件相关电位检测；情绪识别；睡眠分期。在前两个范式上，SSML 取得了超过 15% 的改善，在第三个范式上，SSML 取得了 4.9% 的改善。结果证明，SSML 方法在 BCI 应用中的很有效。

**1014,** 特征归因(feature attribution)方法的有效性及其与自动评估分数的相关性

Giang Nguyen, Daeyoung Kim, Anh Nguyen, 2021

奥本大学、韩国科学技术院

在现实世界的高风险应用中，解释人工智能模型的决策越来越重要。数以百计的论文提出了新的特征归因方法，并在相关的工作中讨论和利用了这些工具。然而，尽管最终的目标用户是人类，大多数的归因方法只在代理自动评估指标(proxy automatic-evaluation metrics)上进行了评估。本文首次对用户进行了研究，衡量了当图像是正常或者是包含对抗性扰动时归因图对协助人类在 ImageNet 分类任务和 Standford Dogs 细粒度分类任务中的有效性。总的来说，让人诧异的是特征归因方法并不比向人类展示最近的训练样本更有效。在 Standford Dogs 细粒度分类这一更困难的任务上，向人类展示归因图并没有帮助，相反，与单独使用人工智能相比，人类-人工智能团队的表现更差了。重要的是，我们发现自动归因图的评估方式与人类-人工智能团队的实际表现相关性很低。我们的发现鼓励社区在人类参与的下游任务中进行更严格的测试，并重新思考现有的评估指标。

**1015,** 利用自我解释能力为深度学习图像分类器提供错误检测

论文作者: MohammadMahdi Karimi, Azin Heidarshenas, William W. Edmonson,  
2022. 10. 15

北卡罗来纳州立农业技术大学、英伟达 (NVIDIA)、苹果公司 (APPLE)、SADAINA  
本文提出了一个可自我解释的深度学习 (SE-DL) 系统，应用于可自我错误检测的图像分类问题。自我错误检测是提高深度学习 (DL) 系统安全运行的关键，特别

是在汽车系统等安全关键应用中。SE-DL 系统同时输出类的预测和对该预测的解释，这为系统如何为人类做出预测提供了解释。此外，作者利用提出的 SE-DL 系统的解释来检测系统潜在的类别预测错误。提出的 SE-DL 系统使用了一组概念来产生解释，这些概念是每个输入图像中与该图像的高层次类别相关的、人类可理解的低层次图像特征。作者提出了一种概念选择方法，对所有的概念进行评分，并根据它们对 SE-DL 系统的错误检测性能的贡献，选择其中的一个子集。最后，作者利用提出的 SE-DL 系统，实行了不同错误检测方案，并将其与没有任何 SE-DL 系统的错误检测方案进行比较。

### 1016, OpenXAI: 实现模型解释的透明评估

Chirag Agarwal、Eshika Saxena、Satyapriya Krishna、Martin Pawelczyk、Nari Johnson、Isha Puri、Marinka Zitnik、Himabindu Lakkaraju, 2022. 6. 22  
哈佛大学、卡耐基梅隆大学

本文介绍了 OpenXAI，它是一个全面的、可扩展的开源框架，用于评估和基准事后解释方法。虽然最近的文献中提出了几种类型的事后解释方法（例如特征属性方法），但很少以有效和透明的方式对这些方法进行系统的基准测试。OpenXAI 由以下关键组件组成：（1）灵活的合成数据生成器、各种现实世界数据集、预先训练的模型和最先进的特征属性方法的集合；（2）用于评估解释方法的忠实性、稳定性和公平性的二十二个定量指标的开源实现；（3）一个公开的可解释人工智能的解释基准排行榜。OpenXAI 易于扩展，因为用户可以轻松地评估自定义解释方法并将其纳入我们的排行榜。总的来说，OpenXAI 提供了一个自动化的端到端通道，不仅简化和标准化了事后解释方法的评估，而且提高了这些方法基准的透明

度和再现性。OpenXAI 数据集和数据加载器、最先进的解释方法和评估指标的实现及排行榜在 <https://openxai.github.io/> 上公开提供。

### 1017, 基于神经拟态技术的边缘计算实现智能人脸识别技术

Jae-Woo Kim, Cosmas Ifeanyi Nwakanma, Dong-Seong Kim and Jae-Min Lee,  
2021. 10. 19

韩国金乌国立理工大学

文章研究了基于神经拟态技术的智能边缘计算。神经拟态技术是一种使用纯硬件方式实现智能系统，而不像传统的通过 CPU 和 GPU 通过软件的方式实现智能系统。本文利用神经元逻辑技术的一种设备 NeuroEdge 介绍了智能边缘计算技术，并通过人脸识别测试验证了其性能。结果表明，使用神经形态技术，比如使用 NM500 芯片，可以节省训练系统所需的时间，并且不会带来有效训练需要大量数据集的负担。

### 1018, 终身可靠性感知神经形态计算案例

Shihao Song、Anup Das Electrical, 2022. 10. 20

美国宾夕法尼亚州费城德雷塞尔大学计算机工程

在非易失性存储器上使用神经形态计算能够让使用脉冲计算和仿生学习算法的机器学习任务的性能提高以及能耗降低。但是操作非易失性存储器的高电压会加速神经元 CMOS 电路的老化。本文评估了 10 种前沿的机器学习应用对于可靠性和性能之间权衡的影响，并在其中考虑了负偏置温度不稳定性 (NBTI) 和时变击穿 (TDDB) 两种故障机制。本文为了提高神经形态电路的终身可靠性，引入了周期

性松弛神经元的方法，即周期性的对神经元进行减压。本文主要的性能衡量指标为脉冲间隔周期(inter-spike interval, ISI)和紊乱脉冲个数(disorder spike count, DSC)，而与神经元周期性松弛的指标主要为减压间隔(tDSI)和减压完成时间(tDSC)。本文的实验主要是将 tDSI 从 10ms 增加到 50ms 并观察各指标的变化。对于可靠性，实验结果有三点：1) NBTI 和 TDDB 的老化会随着 tDSI 的增加而增加。2) tDSI 对于老化的影响有着应用差异性，即对于不同的机器学习应用 tDSI 会有着不同程度的影响。3) 相比于 NBTI, TDDB 对于不同应用的老化增加程度很接近。对于性能，实验结果表示：随着 tDSI 的增加 ISI 失真和 DSC 都会减少。对于热影响，实验结果表示在 325K 和 350K 的温度下的老化会比在 300K 温度下的老化增加 7%和 26%。

## 1019, 具有紧凑神经核心的神经拟态计算平台

Pujun Zhou, Shaogang Hu, 2022. 10. 20

电子科技大学

本文作者提出了一种紧凑的神经元模型，减少了硬件开销。通过使用神经元复用技术和权重聚类算法，构建了一个紧凑、通用的神经拟态计算核心。本文作者通过 588 个 LUT 整合 1K 神经元和 1M 突触，提出了一个大型神经拟态计算系统。这个神经拟态计算平台集成了 64 个神经拟态核心和相关控制组件，包括了 64K 个神经元和 64M 个突触。在这项工作中，一个核心集成了多达  $1024 \times 1024$  个突触，作者采用了重量分担技术，通过 Kmeans 聚类，将  $1024 \times 1024$  权重数据聚类为 16-16 位权重数据，每个突触都使用一个 4 位索引来索引相应的权重参数，该权重共享技术对网络分类的精确度的影响约为 1%，但比传统设计使用 16 位存储空间来

存储索引，大大减少了存储压力，减少了内存开销。本文作者在 Xilinx 的 FPGA-Vertex-6 平台上成功部署了该平台，成功地部署了一个三层脉冲神经网络(SNN)，该网络有 784 个输入，隐藏层中  $2 \times 1024$  个神经元，10 个输出神经元，用其来识别 MNIST 数据集，并成功通过该平台实现了 98.41% 的识别准确率。此外，作者还分析了该神经拟态系统的频率和功耗。ISE 报告显示，基于 ASIC 原型平台的神经拟态计算系统具有 460mW 的动态功率和 105MHz 的最大工作频率。

## 1020, HiSMatch: 基于时序知识图谱推理的历史结构匹配

Zixuan Li、Zhongni Hou、Saiping Guan、Xiaolong Jin、 Weihua Peng、 Long Bai、Yajuan Lyu、Wei Li、Jiafeng Guo、Xueqi Cheng, 2022.10.18

中国科学院大学计算机科学与技术学院、中国科学院网络数据科学与技术重点实验室、百度公司

时序知识图谱 (TKG) 是一组有各自时间点的知识图谱序列，采用 4 个方面 (主语，谓语，宾语，时间) 来描述动态事实。其通过回答诸如 (查询实体，查询关系，查询时间) 的关于未来的问题，TKG 促进了许多现实世界的应用。这实际上是一个基于历史结构的查询与候选实体之间的匹配任务，其中历史结构反映了实体在不同时间的行为趋势。此外，现在的知识图谱提供了所有实体的背景知识，这也有助于匹配。因此，在本文中，我们提出了历史结构匹配模型 (HiSMatch)。它应用两个结构编码方式来捕获查询和候选实体的历史结构中包含的语义信息。另外，它也采用了另外一种编码方式将背景知识集成到模型中。在六个标准数据集上进行的 TKG 实验表明，提出的 HiSMatch 模型有显著改进，与最先进的标准相比，MRR 性能提高了 5.6%

## 1021, 使用增强型知识图谱转换器进行跨域方面术语提取

论文作者: Phillip Howard、Arden Ma、Vasudev Lal、Ana Paula Simoes、Daniel Korat、Oren Pereg、Moshe Wasserblat、Gadi Singer, 2022.10.18

美国英特尔实验室、以色列英特尔实验室

方面术语的提取是文本精细粒度情感分析的关键步骤,当训练核测试数据来自同一域时,这项任务现存的方法有着不俗的表现。但是,当测试数据于训练数据来自不同域时,这些方法的效果急剧下降。为了解决这些方法可扩展性和健壮性不足的问题,我们提出了一种新的构建特定域的包含方面术语标识相关的信息的知识图谱的方法。我们引入了一种从这些知识图谱中抽取信息将其注入到转换模型中的方法,该方法包含两种可选的知识插入机制:通过查询扩充以及控制关注模式。我们使用我们的方法在跨域的方面术语提取中得到了极好的表现,另外我们还调查了存在对转换效果有影响外部知识总量。

## 1022, 学习多图上部分匹配网络的宇宙模型

Zetian Jiang, Tianzhe Wang, Jiaxin Lu, Junchi Yan, 2022年10月19日

上海交通大学

本文考虑两个或多个图的部分匹配的一般设置,即一个图中的所有节点不一定能在另一个图中找到它们的对应关系,反之亦然。本文从宇宙匹配的角度来处理这个无处不在的问题,即每个节点要么被匹配到虚拟宇宙图中的一个锚点,要么被视为一个离群者。这种宇宙匹配方案有几个重要的优点,在现有的基于学习的图形匹配(GM)文献中没有被采用。首先,离群者匹配和离群者检测的微妙逻辑可以被清晰地建模,否则在成对匹配方案中处理起来就不太方便。第二,它实现了

端到端的学习，特别是用于匹配的宇宙级亲和度量学习，以及用于聚集异常点的损失设计。第三，产生的匹配模型可以很容易地处理在线匹配下新到达的图形，甚至是来自训练集不同类别的图形。这是第一个可以同时处理双图匹配、多图匹配、在线匹配和混合图匹配的深度学习网络。

**1023**，通过听觉神经反馈训练启用完全闭锁综合征患者的皮质内信号拼写接口  
Ujwal Chaudhary、Ioannis Vlachos、Jonas B. Zimmermann、Arnau Espinosa、Alessandro Tonin 等 2022 年 3 月 22 日

慕尼黑博根豪森诊所、欧洲大学、布朗大学、图宾根大学等

肌萎缩侧索硬化症 (ALS) 患者随着运动神经元变性的进展，可能会失去所有基于肌肉的通讯途径，最终可能会失去任何通讯手段，即进入完全闭锁状态 (CLIS)。之前的研究可以使处于闭锁状态 (LIS) 即眼球运动或凝视控制完好的患者借助辅助和增强交流的设备进行通信。本文通过将两个 64 微电极阵列植入一名处于完全闭锁状态 ALS 患者的补充和初级运动皮层，患者根据听觉反馈调节神经放电频率，并使用此策略一次选择一个字母，形成单词和短语来传达他的需求并进行沟通。这个案例研究提供了证据，证明即使在完全闭锁的状态下，基于大脑的意识交流也是可能的。

**1024**，从低频和跨频频内脑电图特征可以解码想象语音

Timothée Proix , Jaime Delgado Saa, David Poeppel, Werner K. Doyle, Orrin Devinsky , Luc H. Arnal 等 2022 年 1 月 19 日

脑机接口技术发展至此，对发音讲话的解码已经有了一些研究，但对想象的语言

的解码却非常少。主要是因为发音讲话时，当发同样的音，大脑中相同区域就会被激活，但对于想象语言来说，其神经信号过于微弱且多变。所以很难通过学习算法进行解码。论文找到 13 名植入电极用于癫痫评估的患者，通过让这些患者执行发音和想象的语言任务，最终获得了 3 个皮质电信号 (ECoG) 数据集。推进想象语音解码，必须明确两个基本要点：(i) 哪些大脑区域提供了最好的解码潜力；(ii) 在这些空间中，哪些神经特征(例如信号频率、交叉频率或区域相互作用)信息最丰富。到目前为止，向真正的“想象语音”解码系统迈出更进一步的最有效的方法是基于皮质电信号 (ECoG)，目前只记录了在接受手术前评估为难治性癫痫患者的实验信息。在实验过程中，患者通常被要求大声说话、想象说话或想象听到讲话，同时记录 ECoG 信号。基于最近的语音神经处理理论，作者提取了用于未来脑机接口的一致和特定的神经特征，并评估了它们在发音、语音和语音表示空间中区分语音项目的性能。虽然高频活动为显性语音提供了最佳信号，但低频和高频功率以及局部交叉频率都有助于想象语音解码，特别是在语音和语音空间，即知觉空间。这些结果表明，低频功率和跨频动态包含了想象语音解码的关键信息。

## 1025, 基于优先子空间识别的行为相关神经动力学建模

Omid G. Sani、Bijan Pesaran、Maryam M. Shانهchi, 2021.1.24

南加州大学、纽约大学、南加州大学

当前的神经动力学模型是在不考虑行为信号的情况下学习的，无法分离出与行为相关和不相关神经信号，以解释生物特定行为的测量结果。本文开发了优先子空间识别方法 PSID，是一种对神经活动进行建模，同时对其行为相关的动作进行分

离和排序的机器学习算法。本文的机器学习算法可以帮助我们研究有关大脑功能的基本问题。PSID 通过分离与不同大脑功能相关的动态神经模式，可以找到被其他文献方法遗漏的神经模式，是因为之前的方法在搜索神经模式时只考虑大脑信号，而 PSID 能够同时考虑大脑信号和行为信号，使用了 Pesaran 实验室采集的四个现有数据集，其数据集为不同的手臂挥舞速度及眼球运动状态。PSID 对数据集执行建模的数据显示，行为信号比其他方法所显示的维度要低，此外，PSID 更能预测行为的旋转信号，更准确地学习了每个关节行为信号。PSID 提供了一种通用的新工具来揭示同行为相关的神经信号，能更好地解码大脑信号所代表的行为，发现大脑和行为信号之间的常见模式，以开发更强的脑机接口。

## 1026, 注意力可视化之外的 Transformer 可解释性

Hila Chefer<sup>1</sup>、Shir Gur<sup>1</sup>、Lior Wolf

特拉维夫大学、Facebook AI 研究院

自注意力 (self-attention) 机制，特别是 Transformer，在文本处理领域占据着主导地位，并且在计算机视觉的分类任务中越来越受到欢迎。现有的方法要么依赖于获得注意力图 (attention map)，要么是沿着注意力图 (attention graph) 进行启发式传播，以直观的显示某种图像分类的部分。本文提出了一种新的计算 Transformer 网络相关性的方法。该方法基于深度泰勒分解原理 (Deep Taylor Decomposition principle) 标记局部相关性，然后将这些相关性得分通过 (神经网络的) 层 (layers) 进行传播。这种传播涉及到注意力层 (attention layer) 和跳跃连接 (skip connection)，对现有的方法造成了挑战。我们的解决方法基于一个特定的公式，该公式对跨层保持了总体相关性。我们在最近的 Transformer 网

络可视化和文本分文的问题上进行了基准测试(benchmark), 并且展现了相对于现有可解释性方法的显著优势。

### **1027, 物联网上可解释的人工智能: 概述、最新进展和未来方向**

Senthil Kumar Jagatheesaperumal、 Quoc-Viet Pham、 Rukhsana Ruby、 Zhaohui Yang、 Chunmei Xu、 Zhaoyang Zhang, 2022. 11. 02

可解释人工智能(XAI)正在通过增强最终用户对机器的信任来改变人工智能(AI)应用领域。随着连接设备的数量不断增加, 物联网(IoT)市场需要为终端用户提供可信的服务。然而, 现有的文献仍然缺乏对物联网使用 XAI 的系统和全面调查工作。为了弥补这一不足, 在本文中, 作者讨论了 XAI 框架, 重点是其特点和对物联网的支持。作者说明了物联网应用中广泛使用的 XAI 服务, 如安全增强、医疗物联网(IoMT)、工业物联网(IIoT)和城市物联网(IoCT)。作者还通过适当的例子提出了 XAI 模型在这些应用中对物联网系统的实施选择, 并总结了对未来工作的关键推论。此外, 作者还介绍了边缘 XAI 结构的前沿发展和第六代(6G)通信服务对物联网应用的支持, 以及关键推论。简而言之, 本文构成了第一个关于为未来物联网用例需求而定制的基于 XAI 的框架发展的整体汇编。

### **1028, 可解释人工智能(XAI)方法分类综述**

Timo Speith、 2022. 5. 17

萨尔大学

最近, 与可解释人工智能(XAI)相关的出版物激增, 加高了 XAI 和公众间的壁垒。因此, 介绍 XAI 方法分类的文章和评论似乎是获得该领域概述的一种受欢迎

的方式。基于这一想法，目前有一种产生此类分类法的趋势，导致了几种相互竞争的方法来构建。本文将回顾最近构建 XAI 方法分类法的方法，并讨论与之有关的一般挑战以及它们各自的优势和局限性。本文的综述旨在帮助学者了解当前分类法面临的挑战。正如本文将要讨论的那样，在绘制 XAI 领域的图表时，依靠其发现的方法之一可能是不够的。为了修正这个问题，本文将提出并讨论三种可能的解决方案：一种新的分类法，其中包含了已审查的分类法、一个 XAI 方法数据库以及一个决策树，以帮助选择适合的方法。

**1029**, 有关带有被动选择元件的基于 RRAM 的神经拟态计算阵列的非线性选择效应的调查

Shengyu Bao, Zongwei Wang\*, Yaotian Ling, Zhizhen Yu, Yabo Qin, Yimao Cai\*, Ru Huang, 2021.11.3

北京大学微电子研究所

作者对利用带有非线性选择元件的无源 RRAM 交叉棒阵列进行神经形态计算进行了全面的研究。对权重映射方法，电压施加策略等方法进行了设计考虑。通过使用这些方法，消除了推理过程中，非线性选择特性的功能障碍。此外，作者还评估了具有非线性选择特性的无源交叉棒阵列的电压降效应。最后，作者的研究为设计具有非线性选择特性的无源交叉杆神经形态计算系统提供了深刻的分析。为了理解电压施加策略和阵列规模之间的权衡关系，作者设计了实验。并在给定的电压条件下，得到了阵列电压和尺度设计区域的相关参数。为进一步的优化设计提供了方向。

**1030, STSC-SNN: 应用于脉冲神经网络的具有时间卷积和关注机制的时空突触连接**

Chengting Yu, Zheming Gu, Da Li, Gaoang Wang, Aili Wang 等, 2022.10.11  
浙江大学信息科学与电子工程学院

脉冲神经网络是神经形态计算中的算法模型之一，其具有时间信息处理能力、低功耗和高生物可信性等特点。它可以高效提取时空特征，使其适合于处理事件流。然而，SNN 中现有的突触结构几乎是完全连接或空间 2D 卷积，两者都不能充分提取时间依赖性。本文中，作者从生物突触中获得灵感，并提出了时空突触连接 SNN (STSC-SNN) 模型，以增强突触连接的时空感受野，从而建立跨层的时间依赖性。作者结合了时间卷积和关注机制来实现突触过滤和门控功能。其工作侧重于在 SNN 中使用时间维度操作来拓宽突触的时空感受野，以增强 SNN 的时空特征提取。时态运算可以连接到时间维度，其包含在一个可插拔模块中，称为时空突触连接 (STSC)。STSC 模块被设计为放置在空间操作之前，用于聚集时间信息并扩大时空感受野，同时保持原始空间操作。STSC 模块由两个模块组成：时间反应过滤器 (Temporal Response Filter) 和前馈侧抑制 (Feedforward Lateral Inhibition)，它们分别承载突触模型的过滤和门控机制。实验表明，赋予突触模型时间依赖性可以提高 SNN 在分类任务中的性能。此外，作者研究了不同的时空感受野对性能的影响，并重新评估了 SNN 中的时间模块。作者在神经形态数据集上进行了测试，包括 DVS128 手势 (手势识别)、N-MNIST、CIFAR10-DVS (图像分类) 和 SHD (语音数字识别)。结果表明，该模型在几乎所有数据集上都优于最先进的精度。

### 1031, 基于动态关系置信度的知识图谱重要路径发明

Shanqing Yu、Yijun Wu、Ran Gan、Jiajun Zhou、Ziwan Zheng、Qi Xuan

浙江工业大学网络安全研究所、浙江工业大学信息工程学院、浙江警察学院  
现有的知识图谱大多是不完整的，可以通过一些推理算法进行补充。基于路径特征的推理方法由于具有较强的可解释性，在知识图推理和补全领域得到了广泛的应用。但是，基于路径特征的推理方法仍然存在以下几个方面的问题：路径搜索效率低，稀疏任务的路径不足，一些路径对推理任务没有帮助。针对上述问题，本文提出了一种DC-Path方法，结合动态关系置信度等指标对路径特征进行评价，进而指导路径搜索，最后进行关系推理。实验结果表明，与现有的关系推理算法相比，该方法能够从知识图谱中选择当前推理任务中最具代表性的特征，在当前关系推理任务中取得更好的性能。

### 1032, 在不完全场景中利用常识对目标进行定位

Francesco Giuliari, Geri Skenderi, Marco Cristani, Alessio Del Bue and Yiming Wang, 2022. 11. 1

我们提出了一种端到端的解决方案，以解决部分场景中的对象定位问题，我们的目标是估计仅给出场景的部分 3D 扫描的未知区域中对象的位置。我们提出了一种新颖的场景表示促进几何推理，有向空间常识图 (D-SCG)，一种空间场景图来自常识知识库的概念节点。具体而言，D-SCG 的节点表示场景对象，边缘为它们的相对位置。然后，每个对象节点通过不同的常识关系连接到一组概念节点。使用提出了基于图的场景表示，我们使用图神经网络来估计目标对象的未知位置实现了一种新颖的注意力信息传递机制。网络首先预测目标对象之间的相对位置通

过聚合对象节点和概念节点，学习对象的丰富表示 D-SCG。然后合并这些相对位置以获得最终位置。我们使用 Partial ScanNet 评估我们的方法以 8 倍的训练速度将定位精度提高了 5.9%。

### 1033, CircleSnake: 使用圆表示的实例分割

Ethan H. Nguyen, Haichun Yang, Zuhayr Asad, Ruining Deng, Agnes B. Fogo, Yuankai Huo, 2022. 11. 02

Fogo were with the Department of Pathology

圆表示最近被引入为医学成像 优化表示，实现更有效的实例对象检测 球形医疗物体。凭借其在实例上的卓越性能 检测，将圆表示扩展到实例很有吸引力 医疗对象分割。在这项工作中，我们提出了 CircleSnake，一个简单的 基于球形端到端圆轮廓变形的分割方法 医疗对象。与流行的深蛇方法相比，我们的贡献 是三重：(1) 我们将复杂的边界框替换为八边形轮廓 使用无计算且一致的边界圆到圆进行转换 轮廓适应，用于分割球形医疗物体；(2) 圆圈 表示的自由度 (DoF=2) 与 八边形表示 (DoF=8)，从而产生更稳健的分割 性能和更好的旋转一致性；(3) 据我们所知， 提出的 CircleSnake 方法是第一个端到端的圆表示 具有一致圆检测的深度分段流水线法，圆 轮廓建议和循环卷积。关键创新是整合 将带有圆检测的循环图卷积转换为端到端 实例分割框架，由提议的简单一致 圆等值线表示。肾小球用于评估性能 的基准。从结果来看，CircleSnake 增加了平均值 肾小球检测精度从 0.559 到 0.614。骰子分数增加 从 0.804 到 0.849。

### 1034, 用于单通道脑电图睡眠阶段分类的一种基于注意力的深度学习方

Emadeldeen Eldele, Zhenghua Chen、Chengyu Liu、Cuntai Guan 等, 2021. 4. 28  
南洋科技大学、东南大学

自动睡眠阶段分类是测量睡眠质量的一个重要手段, 本文我们提出了一种新颖的方法 AttnSleep, AttnSleep 是基于注意力的深度学习架构, 调用了对睡眠阶段进行分类单通道脑电信号。AttnSleep 的两个模块分别是基于多分辨率的特征提取模块卷积神经网络 (MRCNN) 和自适应特征重新校准 (AFR)。MRCNN 可以提取低频和高频特性, 而 AFR 能够提高通过建模提取的特征之间的相互依赖。这两个模块之后是时间上下文编码器 (TCE) 模块, TCE 是一种多头注意力机制, 以捕获依据时间提取的特征之间的依赖关系。本文还提出了一个类别感知的成本敏感损失函数来处理数据不平衡的问题。在三个公共数据集上的实验结果表明, AttnSleep 在各种评估矩阵上都优于最先进的方法。此外, 本文还进行了消融研究以显示所提出方法中每个模块的有效性, 也进行了敏感性分析, 以证明 MHA 中磁头数量造成的影响。结果表明, AttnSleep 在不同数量的正面时相当稳定。对于未来, 本文考虑迁移学习和领域适应技术, 使用它们以调整在标记数据集上训练的模型以对其他数据集中的未标记睡眠数据进行分类。

### 1035, 多感官反馈脑机接口在脑卒中康复中的应用: 个案研究

Danut C. Irimia、Woosang Cho、Rupert Ortner、Brendan Z. Allison、Bogdan E. Ignat、Guenter Edlinger、Christoph Guger, 2017. 11. 17

雅西工业大学、泰克医疗工程有限公司、雅西医药大学、泰克医疗工程有限公司  
传统治疗方法无法为瘫痪患者提供一个整合了感觉在内的闭环运动康复。本研究

提出了一种基于运动想象 (Motor Imagery, MI) 的脑机接口 (Brain-Computer Interface, BCI)、功能电刺激 (Functional Electrical Stimulation, FES) 和视觉反馈技术的软硬件平台 recoveriX 系统。是一种用于脑卒中后康复的完整的感知-运动闭环治疗系统。该系统在临床环境中对两名慢性中风患者进行了测试。病人被要求以随机的顺序想象左手或右手的运动。

在这两个 MI 任务中，提供了两种类型的反馈：从显示器的左侧或右侧延伸的一根杆作为视觉反馈和 FES 刺激的被动手张开动作作为本体感受反馈。两种类型的反馈都依赖于使用公共空间模式和线性判别分析分类器获得的 BCI 分类结果。经过 10 个疗程的恢复期训练，1 例患者在麻痹性手腕部分恢复了手腕伸展的控制，另 1 例患者中指活动范围增加了 1 厘米。计划对包含一系列更新的新版的 recoveriX 系统进行一项控制组研究。

### **1036, 基于超高密度脑电图的脑机接口系统的单个手指运动解码**

Hyemin S. Lee<sup>1</sup>、Leonhard Schreiner、Seong-Hyeon Jol、Sebastian Sieghartsleitner、Michael Jordan、Harald Pretl<sup>3</sup>、Christoph Guger、Hyung-Soon Park<sup>1</sup>, 2022.10.9

KAIST, g.tec Medical Engineering GmbH,

脑机接口技术使用户无需移动身体就可以操作外部设备。基于脑电图 (EEG) 的脑接口系统因其时间分辨率高、使用方便、可移植性好等优点而受到广泛的研究。然而，研究高空间分辨率脑电图对解码精确身体动作 (如手指动作) 的影响的研究较少，而手指动作在日常生活活动中是必不可少的。通常脑电图系统的空间传感器分辨率较低，可以通过省去传统的脑电图电极分布标准和普通的安装结构来改

善。在本研究中，作者使用新提出的柔性电极网格直接连接头皮，提供超高密度脑电图。他们利用分布在对侧感觉运动皮层的总共 256 个通道来解码单个手指运动，从而探索了这种新系统的性能。uHD 脑电图电极间距离为 8.6 mm，而传统脑电图电极间距离平均为 60~65 mm。5 名健康的受试者参与了实验，根据视觉提示进行单指伸展，并收到了虚拟形象的反馈。

该研究利用  $\mu$  (8-12 Hz) 和  $\beta$  (13-25 Hz) 波段功率特征进行分类和地形图。使用 MNI-152 模板头生成每个频段的 D ERD/S 激活图。采用线性支持向量机 (SVM) 对手指进行成对分类。地形图显示有规律的、聚焦的线索后激活，尤其是在信号质量最佳的受试者中。受试者的平均分类准确率为 64.8(6.3)%，其中中指和无名指为最高平均准确率为 70.6(9.4)%。需要进一步研究使用实时反馈和运动图像任务的 uHD 脑电图系统，以提高分类性能，并为外部设备的 BCI 手指运动控制奠定基础。

### 1037, 基于神经形态计算技术的情感识别系统设计

Ajay B S, 2022. 11. 10

Intel Technology India Pvt. Ltd

人类的情绪能够由表情所感知，通过深度神经网络实现情绪识别在当今自动化世界中取得了广泛的应用。神经拟态计算(后文用 NC 表示)是一种超低功耗技术，且能够同时执行内存的计算和处理。因此，NC 开发大型数据集驱动的 AI 模型的能力，包括可扩展的并行处理，远远超过常规 cpu。传统上，为了补偿设计成本，对于精度较低的权值和激活函数，更倾向采用二进制形式的量化神经网络。1 位二进制量化神经网络是最紧凑的模型，其中权值和激活函数都用 1 位二进制表示。

因为交叉架构只需要两个可能的突触数组来计算一个神经元，在 NC 架构上的二元神经网络实现有望进一步简化实现。并基于此提出了六种基于 NC 的二元突触网络(NC- bsn)的人脸表情情感识别系统。

作者采用线性离子漂移忆阻模型设计了基于 NC 的 BSN 结构，并在 Synopsys VCS 工具中进行了模拟测试。从训练阶段提取训练后的二进制权值，并在忆阻交叉杆模型中进一步编程，根据训练过的权值将输入激活与 LRS 或 HRS 状态绑定。整个任务训练了 90 张图片（训练集），包括快乐、悲伤、愤怒、厌恶、恐惧和惊讶的状态。此外，还测试了 40 张未使用过的图像，以检测分类器的准确性（测试集）。六种情绪输出结果平均准确率为 69.5%。为情感识别系统设计基于 NC 的二元突触结构的模拟工作，目前还没有尝试过。

### 1038, 使用波分复用的神经拟态计算

Xingyuan Xu, Weiwei Han, Mengxi Tan, Yang Sun, Yang Li, Jiayang Wu, Roberto Morandotti, Arnan Mitchell, Kun Xu, and David J. Moss, 2022.9.7

光学神经网络 (ONN) 是人工神经网络的下一代神经形态加速器，光学神经网络或光学神经形态硬件加速器具有显著提高主流电子处理器的计算能力和能效的潜力。为了实现 ONN，需要形成神经元之间互连的加权突触在时间、空间和波长划分并行性内建立的多个物理路径。

本文回顾了回顾了基于时分、空分和波分复用 (SDM、TDM 和 WDM) 的 ONN 的最新工作，采用了不同的复用技术来演示 ONN，其中波分复用 (WDM) 技术充分利用了光学器件在宽带方面的具有独特优势，重点介绍了使用集成微码实现 ONN 的方法。本文还介绍了基于 WDM 的卷积加速器的原理和工作原理以及人体图像处理的实验

结果。作者还提出，尽管 ONN 取得了显著进展，但还有许多问题亟待解决，如需要实现光子学的密集集成系统、需要在芯片上演示更多计算操作和网络架构以展示通用性、需要定制算法来克服制造缺陷和芯片上串扰的挑战等等。

### 1039, 用于节能神经形态计算的有机光电突触器件

Qingxuan Li , Tianyu Wang , Xuemeng Hu, Xiaohan Wu, Hao Zhu , Member, IEEE, Li Ji , Member, IEEE, Qingqing Sun , David Wei Zhang, and Lin Chen, 2022. 7. 7

有机材料由于良好的生物兼容性和机械灵活性的特点，所以在光电神经形态计算中有很大的应用潜力。这种用于神经形态计算的光电子突触是在柔性基底上制造的，并且它优良的铁电性会赋予器件大于 18V 的存储窗口以及更加稳定的电导调制。并且 3,2-b 和 C8-BTBT 优良的光电特性会让设备在极低的电压下让每个事件拥有极低能耗的工作。而且在光电双重调制下，改突触装置能够模拟类似于兴奋性突触后电流 (EPSC)、长期增强/抑制 (LTP/LTD) 的生物突触行为。本文主要研究了在固定光强下该人工突触对不同波长光脉冲的响应。实验发现该突触装置在 360nm 波长下具有最大的突触电流，并且发现在施加光尖峰的瞬间，突触后电流增加到 4.7nA，然后缓慢衰减。本文还基于该器件优良的光电双调制特性，利用有机铁电突触装置构建了一个人工神经网络，并在 MNIST 数据库中进行手写识别的测试，最后测试的准确率达到 90.6%。本文的研究为开发多功能人工突触装置和人工神经网络系统提供了有效途径。

**1040**, 将 AlphaFold 训练时间从 11 天减少到 67 小时的 FastFold

Shenggan Cheng, Ruidong Wu, Zhongming Yu, Binrui Li, Xiwen Zhang, Jian Peng, Yang You, 2022.3.2

蛋白质结构预测是结构生物学领域理解基因翻译和蛋白质功能的重要方法。现有的 AlphaFold 模型将 Transformer 模型引入了具有原子精度的蛋白质结构预测领域,然而由于 AlphaFold 模型的性能特征和巨大的内存消耗,训练和推理既耗时又昂贵。本文提出了 FastFold 模型, FastFold 是一种用于训练和推理的蛋白质结构预测模型的更高效实现,包括一系列基于对 AlphaFold 性能的全面分析的 GPU 优化,同时通过动态轴向并行和对偶异步操作,实现了模型的并行,大幅地提升了效率,超越了现有流行的模型及其并行技术。实验结果表明, FastFold 将整体训练时间从 AlphaFold 的 11 天减少到 67 小时,并实现了 7.5 至 9.5 倍的长序列推理加速。此外,本文将 FastFold 扩展到 512 个 GPU 上,并以 90.1% 的并行效率实现了  $6.02 \times 10^{15}$  浮点运算次数/秒。

**1041**, 使用 AlphaFold 进行高度准确的蛋白质结构预测

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, 等, 2021.7.15

DeepMind 团队

蛋白质对生命至关重要,了解它们的结构可以促进对其功能的机械理解。通过巨大的实验努力,已经确定了大约 100000 种独特蛋白质的结构,但这只是数十亿已知蛋白质序列中的一小部分。结构覆盖率因确定单个蛋白质结构所需的数月至数年的艰苦努力而受到限制。需要精确的计算方法来解决这一差距,并实现大规模

结构生物信息学。仅仅基于氨基酸序列预测蛋白质将采用的三维结构——“蛋白质折叠问题”的结构预测部分——50 多年来一直是一个重要的开放研究问题。尽管最近取得了进展，但现有的方法远远达不到原子精度，尤其是当没有同源结构可用时。在这里，本文提供了一种计算方法，即使在没有已知类似结构的情况下，也可以以原子精度定期预测蛋白质结构。在具有挑战性的第 14 次蛋白质结构预测关键评估 (CASP14) 中，本文验证了基于神经网络的模型 AlphaFold 的完全重新设计版本，证明了在大多数情况下与实验结构具有竞争力的准确性，并大大优于其他方法。AlphaFold 最新版本的基础是一种新的机器学习方法，它将关于蛋白质结构的物理和生物学知识，利用多序列比对，结合到深度学习算法的设计中。

## 1042, AlphaFold 及其对内在紊乱蛋白质的意义

Kiersten M. Ruff、Rohit V. Pappu, 2021. 10. 1

华盛顿大学

从蛋白质的氨基酸序列精确预测蛋白质的三维结构已经成熟。AlphaFold 是一种基于深度学习的蛋白质结构预测方法，在预测精度的独立评估中显示出显著的成功。结构生物信息学的一个重要时代是人类蛋白质组中超过 98% 的蛋白质序列的结构注释。有趣的是，许多预测具有非常低的置信度区域，并且这些区域与内在无序区域 (IDRs) 有很大的重叠。蛋白质组中超过 30% 的区域是无序的，这与过去 20 年的估计一致，因为已经进行了大量的努力，以推广结构-功能范式，包括构象异质性和动力学的重要性。有了来自 AlphaFold 的结构注释，就容易对 idr 及其交互体的“结构”进行推断。在这里，文章对可能出现的误解提出了警告，并强调了为 idr 序列-集成-功能关系提供具体理解的努力。这一观点旨在强调 idr

在序列-函数关系 (SERs) 中的重要性，并强调如何着手提取定量的 SERs，以理解 idr 如何发挥作用。

#### 1043, 可解释性的概念和可解释人工智能的评价方法

论文作者: Giulia Vilone1、Luca Longo1, 2021

都柏林理工大学

可解释 AI 在过去的几年中经历了大幅的增长。这得益于机器学习的广泛应用，特别是深度学习，导致如今发展的高精度模型都在一定程度上缺乏 explainability 和 interpretability。大量解决这一问题的方法已经被提出、开发和测试，同时还有一些研究试图定义可解释性的概念及其评价方法。本文对知识体系做出了贡献，它通过一个分层系统对所有的科学研究进行了分类，对与可解释性概念和 XAI 方法的评价方法相关的理论和概念进行了分类。这个层次结构建立在对现有分类法和同行评议的科学材料的详尽分析之上。研究结果表明，学者们已经确定了解释应该满足的许多概念和要求，以便让用户容易理解，并提供可操作的信息，为决策提供依据。他们还提出了各种方法来评估机器生成的解释在多大程度上满足这些要求。总的来说，这些方法可以归纳为以人为本的评估和以更客观的指标进行的评估。然而，尽管围绕可解释的概念发展了大量的知识，但学者们对于如何定义解释以及如何评估其有效性和可靠性并没有达成普遍共识。最后，本文通过批判性地讨论这些差距和局限性而得出结论，并定义了未来的研究方向，将可解释性作为任何人工智能系统的起始组成部分。

#### 1044, 一种一致有效的归因方法评价策略

Yao Rong, Tobias Leemann, Vadim Borisov, Gjergji Kasneci, Enkeleajda Kasneci, 2022.11.01

随着近年来提出的各种局部特征归因方法，后续文章提出了几种评估策略。为了评估不同归因技术的归因质量，图像领域中最流行的这些评估策略使用像素扰动。然而，最近的研究发现，不同的评估策略会产生相互冲突的归因方法排名，而且计算起来可能非常昂贵。在这项工作中，作者们提出了基于像素扰动的评估策略的信息论分析。本文的发现表明，与像素的实际值相反，通过去除像素的形状泄漏的信息严重影响了结果。利用本文的理论见解，作者提出了一个新的评估框架，称为移除和 Debias (ROAD)，它提供了两个贡献：第一，它减轻了混杂因素的影响，这需要评估策略之间更高的一致性。其次，ROAD 不需要昂贵的再训练过程，与现有技术相比，可节省 99% 的计算成本。

#### 1045, DeepFace-EMD: 使用 Patch-wise Earth Mover 距离重新排序改进分布外人脸识别

Hai Phan<sup>1</sup>, Anh Nguyen<sup>1</sup>, 2022

奥本大学、卡内基·梅隆大学

人脸识别 (Face Identification, FI) 无处不在，并促使执法部门做出许多高风险的决定。目前最先进的人脸识别方法通过获取两张图像嵌入 (image embedding) 的余弦相似度 (cosine similarity) 来比较两张图像。然而，这种方法对训练集或图库中不包括的新类型的图像 (例如，当查询的人脸被遮盖、裁剪或旋转时)，即分布外图像 (out-of-distribution, OOD) 的泛化能力很差。本文提出了一个重

新排名的方法，使用推土距离 (Earth Mover's Distance, EMD) 来比较两张脸的深层空间特征的图像区别 (patch)。我们的额外比较阶段明确地检查了细粒度的图像相似性 (例如，眼睛和眼睛)，并且比传统的 FI 相比，对 OOD 扰动 (perturbation) 和遮挡 (occlusion) 更鲁棒 (robust)。有趣的是，在不对特征提取器进行微调的情况下，我们的方法在所有测试的 OOD 查询中都能提高准确性：不论是遮蔽的 (masked)、裁剪的 (cropped)、旋转的 (rotated) 和对抗性的 (adversarial)，同时在相似分布 (in-distribution) 的图像上也能获得类似的结果。

**1046**， 将对比学习和知识图嵌入相结合，为意大利语开发医学单词嵌入

Denys Amore Bondarenko、Roger Ferrod、Luigi Di Caro，2022.11.09

单词嵌入在当今的自然语言处理任务和应用中发挥着重要作用。虽然预先训练的模型可以直接使用并集成到现有的管道中，但它们通常会经过调整以更好地适应特定的语言或领域。在本文中，我们试图通过对比学习 (CL) 和知识图嵌入 (KGE) 的结合来改进意大利医学领域未被发现的利基中的可用嵌入。主要目的是提高医学术语之间语义相似性的准确性，这也被用作评估任务。由于意大利语缺乏医学文本和受控词汇，我们通过结合现有的 CL 方法 (多重相似性损失、语境化、动态采样) 和 KGE 的整合，开发了一种特殊的解决方案，创造了一种新的损失变体。尽管没有超过以多语言模型为代表的最新技术，但所获得的结果令人鼓舞，与初始模型相比，性能有了显著的飞跃，同时使用的数据量也大大减少。

**1047**， DeepE: 用于知识图谱嵌入的深度神经网络

朱丹浩、沈思、黄书剑、尹畅、丁子奇, 2022. 11. 09

南京理工大学犯罪科学系、江苏警察学院、南京大学计算机科学系

近年来, 基于神经网络的方法在知识图谱嵌入 (KGE) 任务中显示了其学习表达性更强的能力。然而, 在简单的图谱上, 深层方法的性能往往落后于浅层方法。一个可能的原因是, 深度模型很难训练, 而浅模型可能足以准确表示简单 KG 的结构。在本文中, 我们提出了一个基于神经网络的模型——DeepE 来解决这个问题, 该模型堆叠多个构建块, 根据头部实体和关系来预测尾部实体。每个构件都是一个线性函数和一个非线性函数的相加。堆叠的块相当于一组具有不同非线性深度的学习函数。因此, DeepE 允许深函数学习深特征, 浅函数学习浅特征。通过大量的实验, 我们发现 DeepE 优于其他最先进的基线方法。DeepE 的一个主要优势是健壮性。DeepE 的平均 Rank (MR) 评分比 FB15k-237、WN18RR 和 YAGO3-10 的最佳基线方法分别低 6%、30%、65%。我们的设计使得在 KGE 上训练更深层的网络成为可能, 例如在 FB15k-237 上训练 40 层, 并且不会在简单的关系上破坏精度。

#### 1048, 用三元对比学习缓解开放知识图谱的稀疏性

Qian Li、Shafiq Joty、Daling Wang、Shi Feng、Yifei Zhang, 2022. 11. 8

东北大学、南洋理工大学、Salesforce Research

形式知识的稀疏性和非本体结构的粗糙性使得开放知识图谱 (openkg) 中的稀疏性问题尤为突出。由于链接稀疏, 学习少量实体的有效表示变得困难。我们假设, 通过引入负样本, 对比学习 (CL) 公式在这种情况下可能是有益的。然而, 现有的 CL 方法将 KG 三元组建模为实体的二进制对象, 忽略了关系导向的三元传播模式, 它们太过通用, 即忽略了 openkg 中出现的零次、少次和同构问题。为了解决这

个问题，我们提出了一个基于头、关系和尾之间三元传播模式的三元框架 TernaryCL。TernaryCL 设计了对比实体和对比关系来挖掘具有负实体和关系的三元鉴别特征，引入了对比自我来帮助零和少射实体学习鉴别特征，对比同义词来建模同义实体，对比融合来聚合来自多路径的图特征。大量的基准测试实验证明了 TernaryCL 优于最先进的模型。

#### 1049, 基于域对抗图神经网络的自行车共享需求预测的跨模式知识自适应

Yuebing Liang, Guan Huang, Zhan Zhao, 2022.11.16

对于共享单车系统来说，需求预测对于确保根据预测的需求及时重新平衡可用的单车至关重要。现有的共享单车需求预测方法大多基于其自身的历史需求变化，基本上将其视为一个封闭的系统，忽略了不同交通方式之间的互动。这对共享单车尤其重要，因为它经常被用来补充其他方式的出行（如公共交通）。尽管最近取得了一些进展，但没有任何现有的方法能够利用多种模式的时空信息，并明确考虑它们之间的分布差异，这很容易导致负转移。为了应对这些挑战，本研究提出了一个领域对抗性多关系图神经网络（DA-MRGNN），用于以多模式历史数据作为输入的共享单车需求预测。引入了一个时间对抗性的适应网络，从不同模式的需求模式中提取可共享的特征。为了捕捉跨模式的空间单元之间的相关性，我们调整了一个多关系图神经网络（MRGNN），考虑到跨模式的相似性和差异性。此外，我们还开发了一种可解释的 GNN 技术，以了解我们提出的模型如何进行预测。我们使用纽约市真实世界的共享单车、地铁和叫车数据进行了广泛的实验。实验结果表明，与现有的方法相比，我们提出的方法性能优越，不同的模型组件也很有效。

## 1050, H2-金毛猎犬: 循证氢研究资助的方法和工具

Paul Seurin, Olusola Olabanjo, Joseph Wiggins, Lorien Pratt, Loveneesh Rana, Rozhin Yasaei, Gregory Renard, 2022.11.15

氢气有望在经济去碳化中发挥重要作用。由于需要发现、开发和了解低成本、高性能、耐用的材料,以帮助最大限度地提高电解成本,以及需要一个智能工具来使基于证据的氢气研究资助决策相对更容易,因此我们开发了 H2 Golden Retriever (H2GR) 系统,利用自然语言处理 (NLP)、知识图谱和决策智能来发现和表示氢气知识。该系统代表了一种新颖的方法,囊括了基于证据的研究资助的最先进技术。相关的氢气论文被从网络上刮取并编入索引,并通过去除噪音和停顿词、语言和拼写检查、词干化和外移化来进行预处理。NLP 任务包括使用 Stanford 和 Spacy NER 进行命名实体识别,使用 Latent Dirichlet Allocation 和 TF-IDF 进行主题建模。由于氢气生产领域的本体,知识图谱模块被用于生成有意义的实体及其关系、趋势和相关 H2 论文的模式。决策智能组件为利益相关者提供了一个成本和数量依赖的模拟环境。PageRank 算法被用来对感兴趣的论文进行排名。在提议的 H2GR 上进行了随机搜索,结果包括按相关度得分、实体、实体之间的关系图、H2 生产的本体论和显示组件互动性的因果决策图进行排名的论文清单。专家们进行了定性评估,认为 H2GR 的功能达到了令人满意的水平。

## 1051, 在面向功能对象的搜索算法网络中利用知识检索提取任务树

Tyree Lewis, 2022.11.15

南佛罗里达大学

功能性面向对象网络 (FOON) 已经发展为一种可用于机器人执行任务的知识表示

方法。一个 FOON 可以看作为机器人提供有序计划的图表，通过知识检索处理。我们比较了提取任务树时的两种搜索算法以评估其性能；迭代深化搜索（IDS）和两种不同启发式的贪婪最佳优先搜索（GBFS）。然后，我们确定哪种算法能够使用最少的功能单元的数量获得各种烹饪食谱的任务树。初步结果表明，根据提供给搜索算法的配方，每个算法可以比其他算法执行得更好。

### 1052, 通过修正矢量的公平可解释学习

Mattia Cerrator, Marius K ¨oppel, Alexander Segner, Stefan Kramerl, Johannes Gutenberg

美因茨大学

神经网络结构已被广泛用于公平表征学习 (fair representation learning) 的环境中，其目的是为一个给定的向量学习一个独立于敏感信息的新表征。文献中提出了各种“去表征化” (representation debiasing) 技术。然而，由于神经网络本身是不透明的，这些方法很难理解，这限制了它们的作用。我们提出了一个新的公平表征学习框架，核心是学习“修正向量”，其维度与给定数据向量相同。然后，修正值被简单地加到原始特征上，因此可以被分析为对每个特征的明确惩罚或奖励。我们通过实验表明，以这种方式约束的公平表示学习问题并不影响性能。

### 1053, Seamful Xai: 在可解释的人工智能中实现 Seamful 设计

论文作者: UPOL EHSANQ、VERA LIAO、SAMIR PASSIMARK O. RIEDL、HAL DAUMÉ

III, 2022. 11. 02

美国佐治亚理工学院、加拿大微软研究院、美国微软、美国佐治亚理工学院、美国马里兰大学、美国微软研究院

人工智能系统中的错误是不可避免的，是由技术限制和社会技术差距引起的。虽然人工智能系统的黑盒化可以使用户体验无缝化，但隐藏接缝有可能使用户失去权力，无法减轻人工智能错误带来的后果。虽然可解释人工智能（XAI）主要是解决算法不透明的问题，但作者提出，有缝设计可以通过战略性地揭示社会技术和基础设施的不匹配来促进以人为本的 XAI。作者通过以下方式引入 "有缝 XAI" 的概念：(1) 在概念上将 "有缝" 转移到人工智能背景下；(2) 开发一个设计流程，帮助利益相关者进行有缝设计，从而增强可解释性和用户代理。作者与 43 位人工智能从业者和用户一起探索这个过程，使用基于场景的共同设计活动，并以真实世界的使用案例为依据。作者分享了经验性的见解、影响以及对这一过程如何帮助从业者预测和设计人工智能中的接缝、接缝性如何提高可解释性、赋予终端用户权力以及促进负责任的人工智能的批评性反思。

#### 1054, 通用注意力模型解释双模态和编码器-解码器变换器的可解释性

Hila Chefer<sup>1</sup>、Shir Gur<sup>1</sup>、Lior Wolf, 2021.3.29

特拉维夫大学、脸书 AI 研究院

Transformer 越来越主导多模态推理任务，如视觉问题回答，由于其使用自我注意和注意力机制将信息情境化的能力，他们取得了最先进的结果。这些注意力模块也在其他计算机视觉任务中发挥作用，包括物体检测和图像分割。与仅使用自我注意力的变形金刚不同，具有注意力集中的变形金刚需要并行考虑多个注意力图，以突出与模型输入中的预测相关的信息。在这项工作中，本文提出了第一种

方法来解释任何基于 Transformer 的架构的预测，包括双模 Transformer 和具有共同关注的 Transformer。本文提供通用解决方案，并将其应用于这些架构中最常用的三种：(i) 纯自我关注，(ii) 自我关注与共同关注相结合，以及 (iii) 编码器-解码器关注。结果表明，本文的方法优于所有现有的基于单模态解释的方法。

### 1055, 用于对话中情绪识别的上下文和情感感知网络

Geng Tu、Jintao Wen、Cheng Liu、Dazhi Jiang、Erik Cambria, 2022.2.7

汕头大学、南洋科技大学

对话中的情感识别 (ERC) 在推荐系统、人机交互和医疗保健等许多领域都具有广阔的应用前景。与其他情绪识别任务相比，对话本质上是一个动态交互的过程，在这个过程中，人们经常依赖上下文和常识知识来传达情绪信息。本文提出了一个上下文和情感感知框架 Sentic GAT 以解决这一挑战。Sentic GAT 中的常识知识由基于情感一致性的上下文和情感感知图注意力机制动态表示，上下文信息由具有层次多头注意力 (HMAT) 的对话转换器 (DT) 捕获，其中 HMAT 是用于获取历史话语对自身和其他话语的依赖性，以获得更好的上下文表示。本文探索了一种对比损失来区分情绪识别中的上下文无关和上下文敏感的话语，以增强直接表达想法的直接对话中的上下文表示。实验结果表明，上下文和情感信息可以促进常识知识的表示，上下文话语的内部和相互依赖性有效地提高了 Sentic GAT 的性能，并且使用情绪强度的 Sentic GAT 在测试数据集上的表现优于最先进的模型。

### 1056, 基于不确定性的多模式融合多目标优化有源说话人检测

Baptiste Pouthier、Laurent Pilati、Leela K. Gudupudi、Charles Bouveyron、Frederic Precioso, 2021. 9. 12

NXP 半导体、蔚蓝海岸大学

从各种各样的研究中已经证实, 将视频和音频数据结合起来检测活跃的说话人有很大的好处。然而, 任何一种方式都可能通过诱导不可靠或欺骗性信息而误导视听融合。本文概述了主动说话人检测作为一个多目标的学习问题, 利用最佳的每一个模式使用一种新的自我注意, 基于不确定性的多模态融合方案。结果表明, 所提出的多目标学习体系结构在提高 mAP 和 AUC 分数方面优于传统方法。我们进一步证明, 在主动说话人检测中, 我们的融合策略优于其他学科中报道的模态融合方法。结果表明, 该方法显著提高了 AVA-ActiveSpeaker 数据集的性能。

### 1057, 用 AlphaFold-Multimer 预测蛋白质复合物

Richard Evans<sup>1</sup>、Michael O'Neill<sup>1</sup>、Alexander Pritzel<sup>1</sup>、Natasha Antropova<sup>1</sup>、Andrew Senior<sup>1</sup>、Tim Green<sup>1</sup> 等, 2022. 3. 10

DeepMind, London, UK, These authors contributed equally,

虽然由于最近的 AlphaFold 模型, 绝大多数结构良好的单链蛋白质现在可以被高精度地预测, 但在许多情况下, 多链蛋白质复合体的预测仍然是一个挑战。在这项工作中, 文章证明了一个专门为已知化学计量学的多聚物输入训练的 AlphaFold 模型, 文章称之为 AlphaFold-multimer, 与输入适应的单链 AlphaFold 相比, 显著提高了预测多聚物接口的准确性, 同时保持了高的链内精度。在没有模板的 17 个异二聚体蛋白的基准数据集中([2]中引入), 文章在 13 个靶标上实

现了至少中等精度 ( $\text{DockQ}[3] \geq 0.49$ ), 在 7 个靶标上实现了高精度 ( $\text{DockQ} \geq 0.8$ ), 相比之下, 之前的先进系统(中基于 alphafold 的系统)在 9 个靶标上实现了至少中等精度和 4 个靶标上实现了至少高精度。文章还预测了 4446 个近期蛋白质复合体的大型数据集的结构, 从这些数据中文章对所有具有低模板标识的非冗余接口进行了评分。对于异质体接口, 他们在 70% 的情况下成功预测接口 ( $\text{DockQ} \geq 0.23$ ), 在 26% 的情况下产生高精度预测 ( $\text{DockQ} \geq 0.8$ ), 分别比 AlphaFold[4] 的灵活连接器修改提高了 +27 和 +14 个百分点。对于同体接口, 他们成功预测了 72% 的情况下的接口, 并在 36% 的情况下产生高精度预测, 分别提高了 +8 和 +7 个百分点

## 1058, 适用于人工智能和神经拟态计算的硅光子学

Bhavin J. Shastri、Thomas Ferreira de Lima、Chaoran Huang、Bicky A. Marquez, 2021.11.16

普林斯顿电子工程系

神经形态计算的兴起归因于当前计算机的计算能力和计算需求之间日益扩大的差距。并由此催生了对新颖的大脑启发式算法的研究, 以及配套的神经拟态处理器的开发设计。这些算法试图在消耗较少能量的情况下实时解决人工智能(AI)任务。作者认为, 基于光子传递信息的高并行性和高速性, 将同样的神经拟态算法应用到需要多通道的多千兆赫模拟信号的应用中, 会得到比数字处理高很多的处理速度。通过将光子设备的高带宽和并行性与神经网络模拟获得的适应性和复杂性相结合, 光子神经网络有潜力比最先进的电子处理器至少快 1 万倍, 同时每次计算消耗更少的能量。

神经形态光子方法可分为两大类：相干(单波长)和非相干(多波长)方法。基于储藏池计算的神经形态系统和 Mach-Zehnder 干涉仪使用的是相干方法。在储藏池计算中，预先确定的隐含层随机权重是不能修改的。另一种方法（非相干）使用硅光子学设计完全可编程神经网络，采用所谓的广播和权重协议。在这种构架下，光子神经元输出具有独特波长的光信号。这些被多路复用成一个单一的波导，并广播到所有其他检测器。一对神经元之间的每个连接都由一个微环谐振器(MRR)权重独立配置，当被单个光电探测器检测到时，波分复用(WDM)载波不相互干扰。Dennard 缩放定律激励着人们提出新的计算方案，从生物拟态计算，到量子计算等种种方式，是人们的尝试成果。而其中神经拟态计算结合硅光子学是一个很好的发展方向。作者希望今后在这个方向能够有更多的人加入，有更多的设计和成果产出。

### **1059, 基于循环光谱切片神经网络的高速光子神经拟态计算**

K. Sozosl , A. Bogris、P. Bienstman、G. Sarantoglou、 S. Deligiannidis、  
C. Mesaritakis, 2022.5.29

使用光子硬件是实现最小功耗皮秒级机器学习处理的神经拟态计算的最有前途的途径之一。在本文中，作者提出了一个新的概念：使用循环光谱切片来实现光子循环神经网络和储层计算架构。它是通过放置在环路中的简单光学滤波器实现的，其中每个滤波器处理输入光信号的特定光谱切片。这种在光子域中实现循环神经处理的新方法，称之为循环光谱切片神经网络。该方案的与光学频域中的直接处理相兼容，从而使特定的神经拟态方法适用于超宽带信号的分解和处理。特别是当需要超快处理时，直接在光学系统中扫描和处理宽光谱，该方法可以满足

最小功耗或数据存储的需求。在作者的方案中，突触权重相当于滤波器中心频率和带宽。ROSS-NN（循环光谱切片神经网络）的另一个关键特性是其循环性质，即ROSS-NN 具有通过时间复用增强输出轨迹数量的特性，这相当于每个空间节点的动态行为的时间延展。如果每个组的环路的特征延迟变化，则可以实现更复杂的网络。作者使用其在要求严格的行业相关任务上进行了数值评估，如 100 Gbaud 的高波特率光信号均衡，该方法表现出了突破性的性能，性能超越了最先进的数字处理技术。与最先进的解决方案相比，作者提出的方法覆盖范围扩大了一倍，同时将复杂性和功耗降低了 10 倍。在这方面，ROSS NNs 可以为实现超高效光子硬件加速器铺平道路，该加速器专门用于处理光通信和高速成像应用中的高带宽光信号。

## 1060, 用 1M1S 实现神经形态计算的人工尖峰传入神经元系统

Sheng Li Fang、Chuan Yu Han、Zheng Rong Han、Weihua Liu 等, 2022.5.5

IEEE

当前，基于尖峰神经网络（SNN）的神经形态计算由于其低能耗和与生物神经系统的高度相似性而引起了广泛的研究兴趣。而对于基于尖峰的神经网络，人工尖峰传入神经元（ASAN）系统是神经形态计算系统与环境交互的重要组成部分。本论文就提出了一个采用一个 VO<sub>2</sub>-Mott 忆阻器和一个电阻传感器（1M1S）的新架构的结构简单的 ASAN 系统。Mott 忆阻器具有双向 mott 转变，良好的耐久性和高均匀性。这个系统通过将柔性压力传感器结合到 1M1S 架构中，并通过将压力刺激转换为速率编码的尖峰来实现触觉 ASAN 系统。其中 ASAN 触觉系统使用了 3x3 阵列，可以很好地识别不同的压力刺激模式。该系统强大适应性能够将其通过广泛

使用的电阻传感器将大量环境刺激转换为速率编码的尖峰，然后将其作为基于 SNN 的神经形态计算的输入。本文通过施加“X”、“J”、“T”和“U”不同的刺激模式并在 3×3 ASAN 系统阵列上得出相应的输出峰值，得出的结果发现所提出的 ASAN 系统可以用于环境信号的识别。因此，所提出的 ASAN 系统可以作为将环境信号转换为基于 SNN 的神经形态计算的速率编码峰值的一个有前途的候选。

### 1061, GraphWOZ: 使用会话知识图谱进行对话管理

Nicholas Thomas Walker、Stefan Ultes、Pierre Lison, 2022. 11. 23

挪威计算中心、班贝格大学

我们提出了一种新的对话管理方法，使用对话知识图谱作为对话状态的核心表示。为此，我们引入了一个新的数据集 GraphWOZ，它包含了《绿野仙踪》对话，即人类参与者与扮演接待员的机器人进行交互的对话。与大多数现有的对话管理工作相比，GraphWOZ 依赖于明确表示为动态知识图谱的对话状态，而不是固定的槽集。这个图谱由不同数量的实体（如个人、地点、事件、话语和提及）和它们之间的关系（如群体的一部分或参加某个活动的人）组成。然后，根据新的观察和系统操作，定期更新图谱。GraphWOZ 随详细的手工注释一起发布，这些注释与用户意图、系统响应以及在用户和系统回合中发生的引用关系相关。基于 GraphWOZ，我们给出了两个对话管理任务的实验结果，即对话实体链接和响应排序。对于会话实体链接，我们展示了如何使用基于字符串和基于图的特征组合的神经模型将话语提及与知识图谱中相应的实体连接起来。然后，通过将图谱的相关内容汇总为文本来执行响应排名，该文本与对话历史连接起来，并作为输入对给定对话状态的可能响应进行评分。

## 1062, H2-金毛猎犬: 循证氢研究资助的方法和工具

ue Tan, Yixin Liu, Guodong Long, Jing Jiang, Qinghua Lu, Chengqi Zhang,  
2022. 11. 24

图神经网络 (GNN) 已经显示出其在图数据建模方面的优越性。由于联合学习的优势, 联合图学习 (FGL) 使客户能够以分布式的方式训练强大的 GNN 模型, 而无需分享他们的私人数据。联合系统的一个核心挑战是非 IID 问题, 这个问题也广泛存在于现实世界的图数据中。例如, 客户的本地数据可能来自不同的数据集甚至领域, 例如社交网络和分子, 增加了 FGL 方法捕捉共同分享的知识和学习通用编码器的难度。从现实世界的图数据集中, 我们观察到一些结构属性被不同的领域所共享, 这为 FGL 中共享结构知识提供了巨大的潜力。受此启发, 我们提出了 FedStar, 一个为图间联合学习任务提取和分享共同基础结构信息的 FGL 框架。为了明确地提取结构信息, 而不是将其与节点特征一起编码, 我们定义了结构嵌入, 并用一个独立的结构编码器对其进行编码。然后, 结构编码器在客户之间共享, 而基于特征的知识则以个性化的方式学习, 这使得 FedStar 能够捕获更多基于结构的领域变量信息, 并避免了特征错位问题。我们对跨数据集和跨领域的非 IID FGL 设置进行了广泛的实验, 证明了 FedStar 的优越性。

## 1063, 在面向功能对象的搜索算法网络中利用知识检索提取任务树

Tyree Lewis, 2022. 11. 15

南佛罗里达大学

功能性面向对象网络 (FOON) 已经发展为一种可用于机器人执行任务的知识表示方法。一个 FOON 可以看作为机器人提供有序计划的图表, 通过知识检索处理。

我们比较了提取任务树时的两种搜索算法以评估其性能;迭代深化搜索 (IDS) 和两种不同启发式的贪婪最佳优先搜索 (GBFS)。然后,我们确定哪种算法能够使用最少的功能单元的数量获得各种烹饪食谱的任务树。初步结果表明,根据提供给搜索算法的配方,每个算法可以比其他算法执行得更好。

#### 1064, 控制与分类器正交的方向

Yilun Xu1, Hao He1, Tianxiao Shen, Tommi Jaakkola1, 2022

麻省理工学院

本文提出确认分类器的方向不变性(directions invariant),以便这些方向可以在风格转移(style transfer)等任务中得到控制。虽然当给定的分类器是线性的时候,正交分解(orthogonal decomposition)是可以直接识别的,但我们在非线性情况下正式定义了一个正交的概念。我们还提供了一种令人惊讶的简单方法来构造正交分类器(利用给定分类器以外的方向的分类器)。通过实验,我们提出了控制正交变化很重要的三个用例:风格转移、领域适应和公平性。当领域在多个方面发生变化时,正交分类器能够实现理想的风格转移,通过标签转移改善领域适应性,并减轻作为预测器的不公平性。

#### 1065, 一种一致有效的归因方法评价策略

Yao Rong, Tobias Leemann, Vadim Borisov, Gjergji Kasneci, Enkeleajda Kasneci, 2022.11.01

德国图宾根大学计算机科学系

随着近年来提出的各种局部特征归因方法,后续工作提出了几种评估策略。为了

评估不同归因技术的归因质量，图像领域中最流行的这些评估策略使用像素扰动。然而，最近的研究发现，不同的评估策略会产生相互冲突的归因方法排名，并且计算起来可能会非常昂贵。本文提出了基于像素扰动的评估策略的信息论分析。本文的发现表明，与实际值相比，通过移除像素的形状，结果受到信息泄漏的强烈影响。根据本文提到的理论见解，最终提出了一个名为 Remove and Debias (ROAD) 的新评估框架，该框架提供了两个贡献：第一，它减轻了混杂因素的影响，这需要评估策略之间更高的一致性。第二，ROAD 不需要计算上昂贵的再培训步骤，与最先进的技术相比，节省了 99% 的计算成本。

## 1066, 传统学习模型和深度学习模型对纵向医疗记录的可解释性

Lin Lee Cheong、Tefagabir Meharizghi、Wynona Black、Yang Guang、Weilin Meng , 2022. 11. 23

亚马逊机器学习解决方案实验室、美国默克公司

深度学习的最新进展使人们对在纵向医疗记录上训练深度学习模型以预测一系列医疗事件产生了兴趣，这些模型显示了很高的预测性能。然而，预测性能是必要的，但还不够，还需要模型的解释和推理来说服临床医生的持续使用。对可解释性的严格评估往往是缺失的，因为模型（传统与深度）和各种可解释方法之间的比较还没有被很好地研究过。此外，评估可解释性所需的基本事实可能是非常主观的，取决于临床医生的观点。该文章的工作是首批评估传统 (XGBoost) 和深度学习 (带注意力的 LSTM) 模型之间和内部的可解释性能的工作之一，这些模型在纵向医疗数据的全局和单个预测层面上。作者使用三种流行的方法来比较可解释性：1) SHapley Additive exPlanations (SHAP), 2) Layer-Wise Relevance

Propagation (LRP), 以及 3) Attention (注意力机制)。这些实现方法被应用于具有设计的基础事实的综合生成的数据集和真实世界的医疗保险索赔数据集。文章表明,总的来说,与 XGBoost 相比,带有 SHAP 或 LRP 的 LSTM 在全局和局部层面上都提供了优越的可解释性,而带有点积注意力的 LSTM 未能产生合理的解释。随着医疗数据量的爆炸性增长和深度学习的进展,评估可解释性的需求将对深度学习模型在医疗环境中的成功采用起到关键作用。

**1067**, 将 AlphaFold 训练时间从 11 天减少到 67 小时的 FastFold

Shenggan Cheng, Ruidong Wu, Zhongming Yu, Binrui Li, Xiwen Zhang, Jian Peng, Yang You, 2022. 3. 2

蛋白质结构预测是结构生物学领域理解基因翻译和蛋白质功能的重要方法。现有的 AlphaFold 模型将 Transformer 模型引入了具有原子精度的蛋白质结构预测领域,然而由于 AlphaFold 模型的性能特征和巨大的内存消耗,训练和推理既耗时又昂贵。本文提出了 FastFold 模型, FastFold 是一种用于训练和推理的蛋白质结构预测模型的更高效实现,包括一系列基于对 AlphaFold 性能的全面分析的 GPU 优化,同时通过动态轴向并行和对偶异步操作,实现了模型的并行,大幅地提升了效率,超越了现有流行的模型及其并行技术。实验结果表明, FastFold 将整体训练时间从 AlphaFold 的 11 天减少到 67 小时,并实现了 7.5 至 9.5 倍的长序列推理加速。此外,本文将 FastFold 扩展到 512 个 GPU 上,并以 90.1%的并行效率实现了  $6.02 \times 10^{15}$  浮点运算次数/秒。

## 1068, Transformer 的简单性偏向及其学习稀疏布尔函数的能力

Satwik Bhattamishra、Arkil Patel、Varun Kanade、Phil Blunsom, 2022. 11. 22

牛津大学、麦吉尔大学

尽管 Transformers 在 NLP 任务上取得了广泛的成功，但最近的研究发现，与递归模型相比，它们很难对几种形式语言进行建模。这就提出了一个问题，为什么 Transformers 在实践中表现良好，以及它们是否有任何属性使它们能够比递归模型更好地泛化。在这项工作中，对布尔函数进行了广泛的实证研究，以证明以下内容：(i) 随机变压器相对更偏向于低灵敏度的函数。(ii) 当在布尔函数上训练时，Transformers 和 LSTM 都优先考虑低灵敏度的学习函数，Transformers 最终收敛到较低灵敏度的函数。(iii) 在具有低灵敏度的稀疏布尔函数上，发现 Transformers 即使在存在噪声标签的情况下也几乎完美地泛化，而 LSTM 过拟合并且实现较差的泛化精度。总体而言，本文结果提供了强有力的量化证据，表明 Transformers 和递归模型的归纳偏差存在差异，这可能有助于解释 Transformers 在表达相对有限的情况下的有效泛化性能。

## 1069, AlphaFold 蛋白质结构数据库: 用高精度模型大规模扩展蛋白质序列空间的结构覆盖面

Mihaly Varadi、Stephen Anyango、Mandar Deshpande、Sreenath Nair1、

Cindy Natassial、Galabina Y ordanova1、David Yuan 等, 2022. 1. 7

European Molecular Biology Laboratory、European Bioinformatics Institute, Hinxton, UK、DeepMind, London, UK,

AlphaFold 蛋白质结构数据库(AlphaFold DB, <https://alphafold.ebi.ac.uk>)

是一个开放的、广泛的高精度蛋白质结构预测数据库。在 DeepMind 的 AlphaFold v2.0 的支持下，它实现了已知蛋白质序列空间结构覆盖的前所未有的扩展。AlphaFold DB 提供了对预测原子坐标、每个残差和成对模型置信度估计和预测对齐误差的编程访问和交互式可视化。AlphaFold DB 最初的版本包含了 21 个模型生物蛋白质组中超过 36 万个预测结构，这将很快扩展到覆盖 UniRef90 数据集中的大多数(超过 1 亿个)代表性序列。

### 1070, 面向高能效嵌入式神经拟态计算的大、小型多微脑核的设计

M. Lakshmi Varshika、Adarsha Balaji、Federico Corradi、Anup Das、Jan Stuijt、Francky Catthoor, 2021.11.22

美国德雷塞尔大学电子计算机工程系

随着基于 spiking 的深度学习 (SDCNN) 推理在嵌入式系统中的应用越来越多，这些嵌入式系统往往会集成神经拟态计算加速器，如  $\mu$ Brain，以提高能效。作者提出了一种基于  $\mu$  brain 的可扩展多核神经拟态硬件设计，以加速 spiking 深度卷积神经网络的计算。为了提高能效，微脑核被设计成在神经元和突触能力上异质(即，大核与小核)，它们之间使用并行分段总线互连，这就使得微脑核与传统的基于网格的片上网络(NoC)相比，有更低的延迟和能耗。作者设计了一个名为 SentryOS 的系统软件框架，将 SDCNN 推理应用程序映射到该设计中。SentryOS 由一个编译器和一个运行时管理器组成。编译器利用大  $\mu$ Brain 核和小  $\mu$ Brain 核的内部架构，将 SDCNN 应用程序编译成子网络。运行时管理器将这些子网络调度到核心上，并将它们的执行流水线化，以提高吞吐量。作者使用五个常用的 SDCNN 推理应用程序评估了所提出的大、小、多核神经拟态计算设计和系统软件

框架，结果表明所提出的设计方案降低了能耗(37%到 98%)、延迟(9%到 25%)，并提高了应用程序吞吐量(20%到 36%)。作者表示，SentryOS 可以很容易地扩展部署到其他的神经拟态计算加速器上，如 Loihi 和 dynap。

### 1071, 超高速神经形态计算系统中具有高效训练协议的光子脉冲神经网络

Dafydd Owen-Newns、Joshua Robertson、Matej Hejda、Antonio Hurtado,  
2022. 11. 22

由于光子技术的出现，新型超快、节能和硬件友好的神经拟态（类脑）计算平台具有了巨大的前景。鉴于 VCSEL 已经显示出在超快 GHz 速率下实现神经元光学尖峰响应的能力，其已经用于实现基于尖峰的信息处理系统。在这项工作中，作者报告了基于仅一个垂直腔表面发射激光器（VCSEL）的硬件友好型光子系统的脉冲神经网络（SNN）操作，以及一种新颖的二进制权重“重要性”训练方案，该方案充分利用了 SNN 用于处理输入信息的光尖峰的离散性质。并且使用高度复杂的多变量分类任务（MADELON）对基于 VCSEL 的光子 SNN 的使用传统的最小二乘训练方法和替代的新二进制加权方案进行测试，比较性能。两种训练方法都达到了 >94% 的优异分类精度，在处理时间的一小部分内超过了数据集的基准性能。新提出的训练方案还显著减少了训练集大小要求以及训练节点的数量（小于总网络节点数的 1%）。因此，这种基于 VCSEL 的光子 SNN 与所报道的“重要性”加权方案相结合，使得基于超快速尖峰的光学处理具有高度降低的训练要求和硬件复杂性，可用于未来的神经形态系统和人工智能应用。

## 1072, 用于神经形态计算的基于漂移和扩散忆阻器的突触中 STDP 特性的 SPICE 研究

Suman Hu , Jaehyun Kang , Taeyoon Kim , Suyoun Lee , Jong Keuk Park , Inho Kim, Jaewook Kim 等, 2022. 1. 19

目前, 神经形态硬件由于能够很好的模仿人类大脑的工作原理, 所有是人工智能之中具有巨大潜力的系统。这种新颖的系统主要利用利用神经元尖峰(动作电位)处理信息以及使用类似于尖峰依赖可塑性(STDP)的生物学方法对神经元之间的突触连接进行训练。而忆阻器就是实现这种神经形态硬件的很好的候选者。当前已经提出了两种忆阻器, 分别为扩散型忆阻器和漂移型忆阻器, 这两种忆阻器都是被提出去对 STDP 进行一个很好的模拟。但是这两种忆阻器对于 STDP 的作用是不相同的, 扩散型主要用于跟踪对 STDP 重要的尖峰定时历史, 而漂移型主要用于将权重信息保持在更长的时间尺度上。本文进行的研究就是利用 SPICE 模型, 系统地研究两个串联的忆阻器在这种突触中的 STDP 特性。本文主要进行的实验是在 SPICE 仿真中, 改变器件参数, 比如时间常数( $\tau$ )、阈值以及导通和关断电阻, 并观察对于突触响应的影响。实验中发现 STDP 曲线中的  $\max\_lw$  和  $\max\_lt$ , 对设备参数非常敏感, 并且发现根据参数存在四种不同形状的 STDP 曲线。这说明 STDP 特性强烈依赖于器件参数, 不同的器件参数会形成不同的曲线, 有些曲线甚至修改了最初 STDP 曲线的形状。本文的结果和分析能够很好的用于新兴的基于设备的突触的设计, 并且可以很好地模拟未来神经形态系统的生物学 STDP 特征。

### 1073, 一种用于增强声学信号的基于时间注意力池的循环卷积架构

Tassadaq Hussain、Wei-Chien Wang、Mandar Gogate、Kia Dashtipour、Yu Tsao、Xugang Lu、Adeel Ahs, 2022. 10

从声音观测中去除背景噪声以获得干净的信号是许多实际声学应用的重要研究课题。由于其在函数映射方面的强大模型能力，基于深度神经网络的算法已成功应用于声学应用中的目标信号增强。由于大多数目标信号在短期和长期上下文中都携带以分层结构编码的语义信息，因此噪声可能会不均匀地扭曲此类结构。在大多数基于深度神经网络的算法中，此类局部和全局效应并未在信号增强的建模架构中明确考虑。本文提出了一种时间注意力池化(TAP)机制与传统的卷积递归神经网络(CRNN)模型相结合，称为TAP-CRNN，它明确考虑了声学信号增强(ASE)的全局和局部信息。在TAP-CRNN模型中，本文首先使用卷积层从声信号中提取局部信息，并使用递归神经网络架构来表征时间上下文信息。其次，本文利用一种新颖的注意机制来上下文处理噪声信号的显着区域。本文使用婴儿哭声数据集来评估，实验结果证实了所提出的TAP-CRNN的有效性，并证明所提出的TAP-CRNN可以更有效地减少婴儿哭声信号中的噪声分量，在不同的信噪比下具有看不见的背景噪声水平。

### 1074, 将增广拓扑的神经进化与卷积神经网络相结合

Jan Hohenheim、Mathias Fischler、Sara Zarubica、Jeremy Stucki, 2022. 12. 1  
苏黎世大学、伯明翰大学

当前的深度卷积网络在其拓扑结构上是固定的。本文探索了通过将扩充拓扑的神经进化(NEAT)与卷积神经网络(CNNs)相结合来使卷积拓扑本身成为参数的可

能性，并提出了使用残差网络（ResNets）块的这样一个系统。然后，解释了本文建议的系统如何只能在进行了额外的优化后才能构建，因为遗传算法比每次反向传播的训练要求更高。并且本文解释了大多数这些流行语，并针对现代机器学习中最重要领域提供了一个温和简短的介绍。

## 1075, AlphaFold 和淀粉样蛋白结构

Universitat Autònoma de Barcelona、Institut de Biotecnologia i Biomedicina、Institut de Biotecnologia i Biomedicina

蛋白质聚集是一种普遍现象，在许多科学领域具有重要意义。虽然淀粉样蛋白的形成通常被认为是有害的，但在所有生物领域中都发现了发挥生理作用的功能性淀粉样蛋白。尽管它们具有功能和病理相关性，但参与淀粉样变过程的大多数分子种类的结构细节仍然难以捉摸。在这里，文章探讨了 AlphaFold 这种高度准确的蛋白质结构预测器在蛋白质聚集领域的应用。虽然作者设想直接应用 AlphaFold 来协助设计具有更高溶解度的球状蛋白质以用于生物医学和工业目的，但使用该算法来预测聚集物种的结构似乎远非微不足道。首先，在淀粉样蛋白疾病中，多种淀粉样蛋白多晶型的存在和聚集中间体的异质性挑战了基于序列的预测所固有的“一个序列，一个结构”范式。其次，异常聚集不是正选择压力的主题，排除了使用基于进化的方法，这是 AlphaFold 管道的核心。相反，淀粉样蛋白多态性似乎受到功能性淀粉样蛋白中定义结构-活性关系需求的限制。因此，它们可以为 AlphaFold 在淀粉样蛋白景观中的应用提供一个起点。

## 1076, 基于在线事件驱动神经拟态计算的电力系统扰动分类

Kaveri Mahapatra、Sen Lu、Abhronil Sengupta、Nilanjan Ray, 2021.12.1

IEEE

输电网络扰动事件的在线准确分类是广域监测的重要组成部分。虽然许多传统的机器学习技术在事件分类方面非常成功,但它们依赖于从控制中心的PMU数据中提取信息,并通过CPU/GPU进行处理,这样的方式会导致很高的能耗。为了降低能耗,同时不降低分类的准确性,本片文章提出了一种新的方式——使用事件驱动的神​​经拟态计算构架来实现动力传输系统扰动的分类。提出了一种基于脉冲神经网络(SNN)的计算框架,利用扰动的稀疏性,促进局部事件驱动操作,对传入数据进行无监督学习和推断。该方法首先提取PMU信号的时空信息并将其编码成脉冲序列,然后利用基于SNN的有监督和无监督学习框架实现分类。此外,利用随网络深度增加的动态神经稀疏峰值事件,给出了深度峰值网络在复杂多类事件识别问题中的优势。此外,作者还提出了一种基于QR分解的选择技术,用于识别参与多个干扰事件低秩子空间的信号。在新英格兰-纽约5区16台机器收集的数据上验证了所提方法的性能。这项工作成功地证明了SNN处理PMU信号用于在线事件驱动的扰动分类的能力。SNN可以利用事件的稀疏性模式,获得与ANN相当的分类精度,同时显著降低能量消耗。此外,提出了一种在多个干扰事件场景下对低秩子空间的信号进行威慑挖掘的信号选择技术,其有效性通过使用所选信号重构多个信号的样本得到验证。在新英格兰-纽约系统的16台5区分类测试中,ANN和SNN的性能比较表明,SNN在MAC到AC操作的计算次数减少20%的情况下,能够达到96.8%的扰动分类精度。与标准神经网络相比,在无监督(有监督)学习框架下的事件驱动SNN预计可降低高达83.6%(82.18%)的能耗。此外,我们

在深度 SNN 上的结果还强调了一个事实，即动态稀疏性的好处将在更深层次的网络中得到最好的实现，这反过来又是复杂机器学习任务的必要条件。因此，随着电网扰动分类问题的复杂性和实时运行场景的增加，SNN 的能耗降低程度和功率效益有望进一步提高。

### 1077, 用于材料内神经形态计算的自适应可编程网络

Kilian D. Stenning、Jack C. Gartside、Luca Manneschi、Jake Love 等，  
2022. 11. 11

现代人工智能和机器学习尽管已经有了不错性能，然而，随着网络规模的增长和冯·诺依曼架构的效率较低，能源成本急剧上升。“存储计算”为大型网络提供了一种节能的替代方案，它可以固定随机权重，以进行低能耗的训练。CMOS 难以支持机器学习的大规模并行处理，而物质神经形态计算是一个很有吸引力的解决方案。纳米磁性人工自旋系统是神经形态硬件的理想选择。它们的被动记忆、状态相关动力学和非线性 GHz 自旋波响应提供了强大的计算能力。然而，任何单个物理储层都必须在性能指标（包括非线性和存储容量）之间进行权衡。在本文中，作者介绍了三种人工自旋系统：方形人工自旋冰、方形人工自旋涡旋冰和无序的风车人工自旋涡旋。作者展示了调整系统几何和动力学如何定义计算性能。作者设计吃的网络，其中每个节点都是高维物理储层，实现并行、深层和多层物理神经网络架构。这解决了物理储层性能折衷的问题，允许一小套协同物理系统来处理不同的任务，并提供广泛的可编程计算的不同配置。这些网络在广泛的任务集上胜过任何单个存储。最重要的是，作者提出的结构超越了储层计算，提出了一种可重新配置编程层间网络连接的方法，实现了按需任务优化性能。

**1078, NC 网络：在基于纵横阵列的非易失性存储器架构上使用聚合子网的高效神经形态计算**

Tao Luo 、 Liwei Yang 、 Xuan Wang , Yingnan Cui 等, 2022.9.9

IEEE

目前，神经形态计算芯片由于是新兴的非易失性存储器的纵横阵列组成的，所以在高能效和高吞吐量的卷积神经网络推理引擎的低功耗实现上有着很大的潜力。但是，这种硬件具有明显的设计限制，比如该硬件的扇入和扇出是有限的，并且资源的映射效率较低，而这些设计上的限制就使得该硬件在 CNN 的设计和部署上有着很大的挑战。正是由于存在这种挑战所以用户必须设计具有复杂硬件架构知识的 CNN 模型，更糟的结果可能是不能将模型与具有高分辨率图像输入的 CNN 硬件相匹配。本文中，面对这种问题就提出了使用传统层结构的约束形式聚合子网 NC 网来解决这些问题。文中提出了一种节能的无缓冲器、模数转换器和数模转换器（ADC/DAC）架构，以及一种可扩展的端到端解决方案，该解决方案能够自动满足纵横式架构的硬件约束，并且同时优化了资源的使用。在该解决方案中，对于 CNN 的探索和部署神经形态纵横阵列硬件是从基于 TensorFlow 的前端设计开始，并且自动化设计流程将 NC 网络从 TensorFlow 映射到纵横阵列架构上。论文中使用各种基准测试了模拟器和现场可编程门阵列（FPGA）模拟器的设计，该设计不仅通过了包括 MNIST、SVHN、CIFAR-10 和 CIFAR-100 的基准测试，还在真实世界应用程序进行了测试，即使用高分辨率（ $224 \times 224$ ）图像作为输入的人类检测。文中提出的系统在基于纵横阵列的神经形态硬件上实现了这些基准的最先进精度，该精度超过了 90%，并且与 TrueNorth 相比，它还使峰值核心使用效率提高了 4.25 倍。

## 1079, 基于可解释机器学习和 Kelly 指数的足球比赛结果预测

论文作者: Yiming Ren、Teo Susnjak, 2022. 11. 30

新西兰梅西大学数学与计算科学学院、

在这项工作中, 作者开发了一种机器学习方法来预测足球比赛的结果。这项研究的创新之处在于利用凯利指数首先将比赛分为几类, 每一类表示不同的预测难度。为了确定该方法的有效性, 作者为每一类比赛开发了使用多种算法的分类模型。除此之外, 还设计了一套以前没有探索过的特征, 包括基于 ELO 的变量。

该数据集源自涵盖 2019-2021 赛季的英超联赛比赛数据。研究结果表明, 将预测问题分解成子任务的过程是有效的, 并产生了与先前的工作相竞争的结果, 而基于集合的方法是最有效的。

文章还设计了一个投资策略, 以便通过与博彩公司的赔率进行比较来评估其有效性。文章开发了一种方法, 通过将凯利指数与预测模型的预定可信度阈值相结合, 使风险最小化。实验发现, 当采用保守的方法, 主要关注预测模型显示出高置信度的容易预测的比赛时, 建议的策略可以带来利润。

## 1080, 通过模型转换进行可解释的强化学习

Mira Finkelstein<sup>1</sup>、Lucy Liu、Nitsan Levy Schlot、Yoav Kolubus、David C. Parkes、Jeffrey S. Rosenshein<sup>1</sup>, Sarah Keren, 2022

耶路撒冷希伯来大学、哈佛大学、以色列理工学院

理解强化学习(Reinforcement Learning, RL)代理新出现的行为可能是困难的, 因为这种代理往往是在复杂的环境中高度复杂的决策程序进行训练的。这就产生了各种用于解释 RL 的方法, 旨在调和代理人的行为和观察者预期的行为之

间可能出现的差异。最近的大多数方法要么依赖于领域知识（可能并不总是可用的），要么依赖于对代理的策略的分析，要么依赖于对底层环境的特定元素的分析，通常被建模为马尔可夫决策过程（Markov Decision Process, MDP）。我们的关键主张是，即使底层 MDP 不完全已知（例如，过渡概率没有被准确学习）或没有被代理人维护（即，当使用无模型方法时），它仍然可以被利用来自动生成解释。为此，我们建议使用正式的 MDP 抽象和转换，以前在文献中用于加速寻找最优策略，以自动产生解释。由于这种变换通常是基于环境的符号表示，它们可以代表对预期和实际代理行为之间的差距的有意义的解释。我们正式定义了这个问题，提出了一类可用于解释突发行为的变换，并提出了能够有效搜索解释的方法。我们在一组标准的基准上演示了这种方法。

**1081, 双重问题：**如何不使用蒙面语言模型合成的反事实来解释文本分类器的决策？

Thang M. Pham、Trung Bui、Long Mai、Anh Nguyen, 2022.10.10

奥本大学、Adobe Research, Adobe 实验室

几十种归因方法背后的一个原则是，将输入特征（这里是令牌）移除之前和之后的预测差异作为其归因。一种流行的输入边缘化（IM）方法（Kim et al., 2020）使用 BERT 代替令牌，从而产生更合理的反事实。虽然 Kim 等人（2020）报告了 IM 是有效的，本文发现这一结论并不令人信服，因为他们论文中使用的 Deletion BERT 度量偏向于 IM。重要的是，这种偏差存在于基于删除的度量中，包括插入、充分性和全面性。此外，我们使用 6 个指标和 3 个数据集进行的严格评估没有发现 IM 优于 Leave One Out（LOO）基线的证据。本文发现 IM 不优于 LOO 的两个原

因：(1) 从输入中删除单个单词只会略微降低分类器的准确性；以及(2) 一个高度可预测的词总是被赋予接近零的属性，而不管它对分类器的真正重要性。相比之下，通过 BERT 使本地可解释模型不可知解释 (LIME) 反事实更加自然，在几个 Rem Ove And Retrain (ROAR) 度量下，一致提高了 LIME 的准确性。

## 1082, 图注释生成对抗网络

Yoann Boget, Magda Gregorova, Alexandros Kalousis, 2022. 12. 01

我们考虑建模高维分布和生成具有与图骨架相关的复杂关系特征结构的数据的新示例的问题。我们提出的模型通过将任务分成两个阶段来解决生成受每个数据点的特定图形结构约束的数据特征的问题。首先，它对与给定图的节点相关联的特征的分布进行建模，第二，它对节点特征有条件地补充边缘特征。我们遵循通过生成对抗网络 (GAN) 结合在节点和边缘集合上操作的置换等变消息传递架构的隐式分布建模策略。这使得能够一次性 (在两个阶段中) 生成所有图形对象的特征向量，并且通过对图表示中的特定节点排序不敏感而有效地使用网络容量。据我们所知，这是第一种沿着图骨架建模特征分布的方法，允许生成具有用户指定结构的注释图。我们的实验证明了我们的模型通过对三个带注释的图形数据集进行定量评估来学习复杂结构分布的能力。

## 1083, 逻辑和常识引导的时间知识图谱

Guanglin Niu, Bo LI, 2022. 11. 30

北京航空航天大学人工智能学院

时间知识图谱 (TKG) 存储的是从涉及时间的数据中导出的事件。由于事件的时间

敏感性，预测事件是极具挑战性的。此外，以往的 TKG 完成度 (TKGC) 方法不能同时表示事件的时效性和因果性。为了应对这些挑战，我们提出了一个逻辑和常识引导嵌入模型 (LCGE)，以共同学习涉及事件时效性和因果关系的时间敏感，以及从常识角度学习事件的时间无关来表示。具体而言，我们设计了一种时态规则学习算法，构建了一种规则引导的谓词嵌入正则化策略，用于学习事件之间的因果关系。此外，我们可以通过辅助常识知识准确地评估事件的合理性。TKGC 任务的实验结果表明，与现有方法相比，我们的模型在性能上有显著的提高。更有趣的是，我们的模型能够从因果推理的角度提供预测结果的可解释性。

#### **1084, 回归校准中余变量的因果选择，以检测连续暴露错误**

Wenze Tang、Donna Spiegelman、Xiaomei Liao、Molin Wang, 2022.11.24

回归校正被用来纠正连续暴露中测量误差引起的因果效应的偏差，但对于如何调整回归校正方法所使用的测量误差模型 (MEM) 和结果模型中的协变量，目前还没有系统的讨论。在本文中，我们研究了 Rosner, Spiegelman 和 Willet (RSW, 1990) 在因果推理框架下为回归校准方法选择最小和最有效的协变量调整集的方法。我们利用有向无环图来说明主题知识是如何帮助选择这种调整集的。我们表明，为了无偏地纠正测量误差，研究者需要在主要研究样本和验证研究样本中收集关于真实暴露和结果的任何共同原因 (1) 以及测量误差和结果的数据。研究人员需要在 MEM 和结果模型中调整这个最小的协变量集。当这些变量只在主要研究中可用时，研究者仍应在结果模型中对其进行调整以减少偏倚，前提是这些协变量与测量误差的关联性较弱。我们还表明，在结果模型中只调整独立于真实暴露和测量误差的结果风险因素可能会提高效率，而调整与真实暴露相关的任何非风险因素

在现实环境中几乎总是导致效率损失。最后，我们将最初提出的估计方法扩展到非参数环境中，其中允许协变量对效果进行修正。

### 1085, 教我讲解：可解释 NLP 数据集的回顾

Sarah Wiegrefel, Ana Marasović<sup>2</sup>, 2021

佐治亚理工学院、华盛顿大学

可解释的自然语言处理 (Explainable NLP, ExNLP) 越来越注重于收集人类标注的解释。这些解释以三种方式被用于下游：作为数据增强以提高预测任务的性能，作为监督以训练模型为其预测产生解释，以及作为基础事实以评估模型产生的解释。在这篇综述中，我们确定了 65 个具有三类主要文本解释的数据集 (亮点、自由文本和结构化)，组织了关于注释每种类型的文献，确定了现有收集方法的优点和缺点，并为将来收集 ExNLP 数据集提出了建议。

### 1086, 用有影响力的例子解释知识图嵌入模型中的链接预测

Adrianna Janik, Luca Costabello, 2022.12.5

埃森哲实验室

作者研究了知识图谱嵌入 (KGE) 模型中解释链接预测的问题。作者提出了一种基于例子的方法，利用知识图谱中节点和边的潜在空间表示来解释预测。作者通过观察去除有影响的三元组后模型性能的逐渐下降来评估识别的三元组的重要性。实验表明，这种产生解释的方法在两个公开可用的数据集的 KGE 模型上优于基线模型 (baseline)。

**1087**, 基于视觉对应的解释提高了人工智能的鲁棒性和人工智能团队的准确性

Giang Nguyen, Mohammad Reza Taesiri, Anh Nguyen, 2022.10.4

Auburn University, 奥本大学

在人类是最终决策者的许多高风险应用中,解释人工智能(AI)预测越来越重要,甚至势在必行。在这项工作中,我们提出了两种新的可自我解释的图像分类器架构,它们首先解释,然后通过利用查询图像和示例之间的视觉对应关系进行预测(而不是事后解释)。我们的模型在分布外(OOD)数据集上持续改进(+1到+4分),而在分布内测试上的表现比 ResNet-50 和 k 近邻分类器(kNN)稍差(-1到-2分)。通过对 ImageNet 和 CUB 的大规模人类研究,我们发现基于通信的解释比 kNN 解释对用户更有用。我们的解释帮助用户比所有其他测试方法更准确地拒绝人工智能的错误决策。有趣的是,我们首次表明,在 ImageNet 和 CUB 图像分类任务中,有可能实现互补的人工智能团队精度(即,高于单独的人工智能或单独的人工)。

**1088**, 神经拟态计算中温度对模拟忆阻器的影响

Yifu Huang、Reed Hopkins、David Janosky、Jack C. Lee 等, 2021.12.8

IEEE

在神经拟态计算中实现的电子设备需要具备可扩展性,以降低制造成本、提高速度和降低功耗。忆阻器因为其能够编码的具有多个模拟状态的偏置历史(biasing history),成为了神经拟态计算设备的候选组件,并展示了内存计算的潜力。忆阻器的线性多电平电导状态可以简化矢量矩阵乘法的实现,对应神经拟态计算器件中的并行处理。

通过精确控制忆阻器中金属丝的延伸和断裂,作者展示了忆阻器件中潜在的多能

级态。研究了温度对忆阻器器件特性的影响。基于实验结果，设计了具有自定义随机梯度下降 (SGD) 优化器的人工神经网络 (ANN) 模型，以模拟训练过程中温度变化对精度的影响以及温度偏差干扰对预训练网络的影响。同时作者对载流子传输行为，新兴记忆技术，以及神经形态计算应用的潜在降温解决方案进行了讨论。这项工作为温度对器件级特性的影响提供了见解，并为潜在的基于忆阻器的神经形态计算应用的电路级设计提供了指导。

最终作者对新兴记忆技术的准确性进行了评估，结果表明，在较低温度下训练的神经网络比在较高温度下训练的准确率更高，功耗更低。预训练网络中的权重值需要使用温度灵敏度的实验数据进行更新。此外，温度干扰会强烈地影响精度。最后，这项基于活化能的设计指导为潜在的基于硬件的神经拟态计算应用提供了额外的见解。

### **1089, 用虚拟神经元在神经形态计算机上编码整数和有理数**

Prasanna Date, Shruti Kulkarni, Aaron Young, Catherine Schuman, Thomas Potok, Jeffrey Vetter, 2022. 8. 15

神经形态计算机通过模拟人脑进行计算，并且具有极低的功率，它们可能会是未来节能计算不可或缺的工具。虽然神经形态计算的传统应用主要依赖于基于 SNN 的机器学习，但最近在图形算法、自主竞赛和线性代数中的应用表明，神经形态计算不仅仅能够实现 SNN。神经形态计算已经证明是图灵完备的，因此能够进行所有通用计算。于是乎，为了充分发挥它在通用、节能计算方面的潜力，设计有效的数字编码机制非常重要。目前在神经形态计算机上编码数字的机制仅限于布尔数或自然数。但是，即使是这些机制也仅限于其开发的特定应用，不适用于通

用计算。此外，其中一些方法由于过度离散化和/或不保留加法而导致数据丢失。在本文中，作者提出了虚拟神经元作为编码正整数、负整数和有理数的机制。作者在 NEST 模拟器中实现了虚拟神经元，并在 8、16 和 32 位有理数上对其进行了测试。作者从理论上比较了虚拟神经元与其他神经形态编码机制的计算复杂性。接下来，作者在神经形态硬件上测试了虚拟神经元，并表明它可以使用基于混合信号忆阻器的神经形态处理器平均使用 23 nJ 的能量执行加法运算。最后，作者通过在五个应用中使用虚拟神经元，证明了虚拟神经元的可用性。此外，作者除了能够证明虚拟神经元是编码有理数的有效机制，还表明虚拟神经元可以模拟具有身份激活功能的人工神经元。

## **1090, 有关内存中的神经形态计算：架构、模拟器和安全性的调查**

Felix Staudigl, Farhad Merchant, 2022.9.9

RWTH Aachen University

这篇论文主要是对内存中神经形态计算进行了调查，目前现有的调查大多致力于硬件或应用层，但是本文作者从一个系统角度来进行分析，包括架构、模拟器和安全性三方面进行调查。架构方面：由于生物大脑及其并行处理大量数据的能力依赖于大量的生物过程，而神经形态计算架构相比于传统架构的优越性就是在于对于这些过程的模仿。但是由于基础技术的不成熟限制了其可行性，因此有了神经启发架构，该架构部分模仿了生物大脑的功能，而非全部。模拟器方面：模拟器是从低级设备建模到高级系统探索的重要工具，模拟器主要分为四个抽象执行模拟的级别，分为系统模拟、架构模拟、电路模拟和设备模拟。通常模拟的时候必须考虑模拟速度和模拟结果精度之间的折衷。系统级别的模拟由于操作的大量性

因此需要高性能，而这样的后果就是精度较低。而设备级别的模拟由于需要较高的计算精度，所以会降低模拟速度。安全性方面：由于 Dennard 缩放定律的崩溃和摩尔定律的放缓，下一代神经形态系统有望成为多种技术的异构集成。而这些系统被认为是具有高性能、低能耗和隐私保护的智能系统。虽然目前对兼容性问题已经展开研究，但是对于安全性问题还没有被发现，而不断演变的威胁将对神经形态平台的安全构成重大挑战，因为攻击向量是无法预料的。

### **1091, 用有影响力的例子解释知识图谱嵌入模型中的链接预测**

论文作者: Adrianna Janik, Luca Costabello, 2022.12.5

埃森哲实验室

研究了知识图谱嵌入 (KGE) 模型中链接预测的解释问题。我们提出了一种基于实例的方法，利用知识图谱中节点和边的潜在空间的表示来解释预测。我们通过观察重要的三元组去除后，模型性能的逐步退化来评估其重要性。我们的实验证明，对于两个公开可用的数据集，这种生成解释的方法优于 KGE 模型的基线。

### **1092, T-STAR: 使用 AMR 图作为中间表示的真实风格转换**

Anubhav Jangra Preksha Nema Aravindan Raghuvver, 2022.12.3

用于训练文本风格转移 (TST) 模型的并行语料库的不可用性是一个非常具有挑战性但很常见的场景。此外，TST 模型隐含地需要在将源语句转换为目标样式时保留内容。为了解决这些问题，通常会构建一个没有风格的中间表示，同时仍然保留源句子的含义。在这项工作中，我们研究了抽象意义表示 (AMR) 图作为中间风格不可知表示的有用性。我们假设像 AMR 这样的语义符号是中间表示的自然选

择。因此，我们提出了 T-STAR：一个由两个组件组成的模型，文本到 AMR 编码器和 AMR 到文本解码器。我们提出了一些建模改进，以增强生成的 AMR 的风格不可知性。据我们所知，T-STAR 是第一个使用 AMR 作为 TST 中间表示的工作。通过彻底的实验评估，我们发现 T-STAR 显著优于现有技术，平均提高了 15.2% 的内容保存率，损失可忽略不计 (~3%)。通过详细的人类评估 (评分为 90000)，我们还发现 TSTAR 的幻觉比最先进的 TST 模型少 50%。

### **1093, 多关系图表示的伪黎曼嵌入模型**

Saeed Paliwal, Angus Brayne, Benedek Fabian 等, 2022.12.08

在本文中，我们将单关系的伪黎曼图嵌入模型概括为多关系网络，并表明将关系编码为流形变换的典型方法从黎曼尼转化为伪黎曼尼的情况。此外，我们构建了一个关系作为多时间流形的独立时空子流形的观点，并考虑了伪黎曼嵌入模型和其威克-旋转黎曼对应模型之间的插值。我们在链接预测的任务中验证了这些扩展，重点是平坦的洛伦兹流形，并证明了它们在生物领域的知识图谱完成和知识发现中的应用。

### **1094, ArcText: 描述卷积神经网络架构的统一文本方法**

Yanan Sun, Gary G. Yen, Bing Xue, Mengjie Zhang, Jiancheng Lv, 2021.11.17

卷积神经网络 (CNN) 的优越性在很大程度上依赖于它们的架构，这些架构通常是用大量的人类专业知识人工提取的，而这种领域知识不一定为每个感兴趣的用户所拥有。现有 CNN 的数据挖掘可以从其架构中发现有用的模式和基本评论，为研究人员提供强大的先验知识，以在他们没有 CNN 专业知识的情况下设计有效的 CNN 架构。现有关于各种最先进的数据挖掘算法的研究，而挖掘 CNN 架构的研究

很少，主要原因之一是 CNN 架构和数据挖掘算法之间的差距。具体来说，当前的 CNN 架构描述无法精确矢量化以提供给数据挖掘算法。本文提出了一种名为 ArcText 的统一方法来描述基于文本的 CNN 架构。特别是在 ArcText 中精心设计了四种不同的单元和一种排序方法，以用足够的信息唯一地描述相同的 CNN 架构。此外，结果描述可以准确地转换回相应的 CNN 架构。ArcText 弥合了 CNN 架构和数据挖掘研究人员之间的鸿沟，并可以被用于更广泛的场景。

### 1095, 非参数掩码语言建模

Sewon Min、Weijia Shi、Mike Lewis、Luke Zettlemoyer, 2022.12.2

华盛顿大学、Meta AI、艾伦人工智能研究所

现有的语言模型 (LM) 在有限的词汇表上用 softmax 来预测标记，这可能使得难以预测稀有标记或短语。本文引入了 NPM，这是第一个非参数掩蔽语言模型，它用参考语料库中每个短语的非参数分布代替了 softmax。实验结果表明，NPM 可以有效地训练出具有对比目标和批量近似的全语料库检索。对 9 个闭集任务和 7 个开集任务的 Zero-shot 评估表明，无论是否使用检索和生成方法，NPM 都显著优于更大的参数模型。在处理罕见模式（词义或事实）和预测罕见或几乎看不见的单词（例如，非拉丁文字）。

### 1096, 在 CASP14 上应用和改进 Alpha Fold

John Jumper、Richard Evans、Alexander Pritzel、Tim Green 等, 2021.10.2

DeepMind, London, UK、Seoul National University, Seoul

文章描述了 Alpha Fold 的操作和改进，该系统由 AlphaFold2 团队进入第 14

届蛋白质结构预测关键评估 (CASP14) 的“人类”类别。CASP14 中输入的 Alpha Fold 系统与 CASP13 中输入的完全不同。它使用了一种新型的端到端深度神经网络, 经过训练可以从氨基酸序列、多序列比对和同源蛋白质中生成蛋白质结构。在评估员的总 z 分数 (2.0) 排名中, Alpha Fold 得分为 244.0, 而次优组的得分为 90.8。Alpha Fold 做出的预测中域 GDT\_TS 为 92.4; 这是 CASP 期间首次达到这种平均准确度水平, 尤其是在更困难的自由建模目标上, 代表了蛋白质结构预测领域的重大进步。他们报告了 Alpha Fold 在 CASP14 期间如何作为一个人类团队运行并进行了改进, 使其现在无需干预即可达到同等水平的性能, 从而为高精度的大规模结构预测打开了大门。

**1097, Alpha Fold 预示着生物学和医学领域的驱动革命**

Janet M. Thornton、Roman A. Laskowski、Neera Borkakoti, 2021.10.12

European Bioinformatics Institute,

使用人工智能预测的蛋白质结构将有助于医学研究, 但如果临床数据可以类似地用于更好地了解人类疾病, 那么最大的好处就会到来。

蛋白质结构预测问题是蛋白质的氨基酸序列如何导致其完全折叠的三维结构的问题。几十年来, 这一直是一项艰巨的计算挑战。

2020 年底, 总部位于伦敦的人工智能 (AI) 公司 DeepMind 宣布取得重大进展, 该公司现为谷歌母公司 Alphabet Inc. 的一部分。DeepMind 的 Alpha Fold 2 程序在两年一度的蛋白质关键评估中明显优于其他方法结构预测 (CASP), 生成质量接近实验测定质量的模型。Alpha Fold 2 已经发布, 最近公开了源代码和来自包括人类在内的各种物种的近 350,000 个蛋白质模型。这一蛋白质结构宝

库对实验和计算结构生物学都有影响，但在这里我们考虑它可能对医学的影响。Alpha Fold 2 使用由结构生物学家收集并由全球蛋白质数据库 (wwPDB) 公开提供的数据，该数据库目前拥有超过 180,000 个实验确定的结构。难能可贵的是，DeepMind 已经把代码和预测放出来供大家使用。

欧洲分子生物学实验室-欧洲生物信息学研究所 (EMBL-EBI) 的 Alpha Fold 蛋白质结构数据库提供了超过 350,000 种蛋白质模型，并提供了查看和查询结构的工具<sup>3</sup>。这些蛋白质来自 21 个物种，包括最常见的模式生物和一些著名的病原体——婴儿利什曼原虫、结核分枝杆菌、恶性疟原虫和克氏锥虫。在今年年底之前，DeepMind 预计将发布涵盖 UniRef90 的模型，UniRef90 是所有已知蛋白质序列的独特样本，包含 1.3 亿种蛋白质。

尽管蛋白质结构本身不会产生新药，但它们通常可以让人们更好地了解蛋白质的分子机制，从而深入了解蛋白质的工作原理以及其调节如何导致疾病或治疗。在过去的 50 年里，蛋白质结构一直是药物设计工作中不可或缺的一部分，许多大型制药公司都建立了自己的结构生物学团队。结构数据在确定给定蛋白质靶点的成药性以及随后能够设计与其结合的小分子药物方面都发挥了关键作用。

## 1098, 别管图表：解决过度挤压而不重新布线的问题

Domenico Tortorella, Alessio Micheli, 2022.12.14

比萨大学

最近的工作已经研究了图瓶颈在阻止消息传递图神经网络中的长距离信息传播中的作用，导致了所谓的'over-squashing'现象。作为补救，已经提出了图重布线机制作为预处理步骤。图回声状态网络 (GESNs) 是一种图的水库计算模型，其

中节点嵌入由一个未经训练的消息传递函数递归计算。在这篇文章中，证明了GESN在不改变图的连通性的情况下，可以在六个异质节点分类任务上获得显著更好的准确率，从而为解决过度挤压问题提供了一种不同的途径。

**1099**，通过面部对齐、训练优化和调度改进基于面部视频的抑郁估计

论文作者：Manuel Lage Cañellas, Constantino Álvarez Casado, Le Nguyen, Miguel Bordallo López, 2022. 12. 13

深度学习模型在使用基于视频的面部表情识别抑郁状态方面显示出可喜的结果，虽然成功的模型通常利用 3D-CNN 或视频蒸馏技术，但在实验中对预训练、数据增强、预处理和优化技术的不同使用使得难以进行公平的架构比较。本文建议通过使用两种特定的面部对齐方法和改进的数据增强、优化和调度技术来增强基于 ResNet-50 的两个仅使用静态空间信息的简单模型。本文在基准数据集上进行的广泛实验获得了与复杂的单流时空模型相似的结果，而两个不同流的分数级融合优于最先进的方法。本文的研究结果表明，预处理和训练过程中的特定修改会导致模型性能的显著差异，并且可能隐藏最初归因于使用不同神经网络架构的实际情况。

**1100**，相关性 CAM：你的模型已经知道该看哪里了

Jeong Ryong Lee、Sewon Kim、Inyong Park、Taejoon Eo、Dosik Hwang, 2022. 12. 8

随着神经网络应用领域的增加和神经网络的发展，解释深度学习模型的能力也变得越来越重要。特别是，在实际应用之前，分析模型的推理和生成结果的过程至关重要。一种常见的解释方法是基于类激活映射（CAM）的方法，通常用于理解计

计算机视觉领域中流行的卷积神经网络的最后一层。在本文中，我们提出了一种新的 CAM 方法，即相关性加权类激活映射（RelevanceCAM），该方法利用分层相关性传播来获得权重分量。这使得解释图对于破碎的梯度问题是可靠和鲁棒的，这是基于梯度的 CAM 方法的共同问题，导致中间层的噪声显著性图。因此，本文提出的方法可以通过正确分析中间层以及最后的卷积层来更好地解释模型。本文可视化了流行图像处理模型的每一层如何使用相关性 CAM 提取类别特定特征，评估了定位能力，并通过实验加权组件证明了为什么基于梯度的 CAM 不能用于解释中间层。在任何深度的层中，相关 CAM 在识别和定位评估方面优于其他基于 CAM 的方法。

## 1101, 语义网络解释

Pei Guol, Ryan Farrell1, 2022

Brigham Young University(杨百翰大学)

网络解释作为揭示网络所学特征的一种努力，在很大程度上仍然是基于可视化的。在本文中，我们的目标是解决过滤器和决策层面的语义网络解释。对于过滤器层面的解释，我们用视觉属性的概率分布表示过滤器编码的概念。决策层面的解释是通过文本总结来实现的，它产生了一个包含网络决策背后线索的解释性句子。我们提出了一种贝叶斯推理算法来自动将过滤器和网络决策与视觉属性联系起来。人类研究证实，语义解释是对可视化方法的有益替代或补充。我们证明了语义网络解释在理解网络的故障模式方面可以发挥的关键作用。更重要的是，语义网络解释能够更好地理解一个模型的性能和它的分布指标之间的相关性，如过滤器的选择性和概念稀疏度。

## 1102, 一种用于社会网络错误信息检测的情态级可解释框架

V ítor Lourenço、Aline Paes, 2022.12.8

弗鲁米嫩塞联邦大学, 戴尔科技

虚假信息的泛滥在全球范围内引起了越来越多的关注, 并产生了重要的社会影响, 激发了事实核查组织的出现, 以减少错误信息的传播。然而, 由人类驱动的核查导致了一项耗时的任务, 并且在它们出现的同时检查可信的信息也是一个瓶颈。由于错误信息不仅与内容本身有关, 也与其他社会特征有关, 本文从多模态的角度探讨了社会网络中的自动错误信息检查。此外, 由于简单地将一则新闻命名为不正确, 可能无法说服公民, 甚至更糟糕的是, 会加强确认偏见, 因此, 本文提出了一个模态级的可解释错误信息分类器框架。文章提出的框架包括一个错误信息分类器, 在可解释方法的辅助下, 产生面向模式的可解释推论。初步研究结果表明, 错误信息分类器确实受益于多模态信息编码, 并且面向模态的可解释机制增加了推断的可解释性和完整性。

## 1103, 神经拟态计算通信网络上异构尖峰神经网络创建的带宽负载和延迟估计网络模拟器

R. Kleijnen、M. Robens、M. Schiek、S. van Waasen, 2021.12.15

Central Institute of Engineering, Electronics and Analytics -  
Electronic Systems (ZEA-2), Forschungszentrum Julich GmbH, Julich,  
Germany

持续几周、几个月或几年, 观察人体内由人脑神经元活动引起的长期学习效果是不切实际的。在过去的十年中, 神经形态计算硬件领域有了显著的发展, 包括

Spinnaker, BrainScaleS 和 Neurogrid。这些新颖的多核模拟平台提供了一种实际的替代方案，通过高水平的细节模拟加速研究大脑中的神经元行为。然而，到目前为止，它们还没有达到人脑的水平，特别是大量的 spike 通信成为了一个瓶颈。在本文中，作者介绍了专门为分析神经形态计算通信网络中不同网络拓扑和通信协议的带宽负载和延迟而开发的网络模拟器(网络模拟器)。与最先进的网络模型和模拟器相比，这个模拟器的独特之处在于，它能够通过不同的模型模拟异构神经连接的影响，以及神经元映射算法的评估。通过将使用同构神经网络的运行结果与可比工作产生的带宽负载进行比较，作者交叉检查了模拟器，但同时显示了我们的模拟器达到的更高的细节水平。通过模拟器，我们展示了基本映射方法对网络带宽负载的影响。这项研究的潜在扩展，重点是开发和评估更有效(和灵活)的映射解决方案，可以使用模拟器来评估映射算法的性能。未来的工作将侧重于对不同的网络拓扑、通信协议和神经网络进行定性和定量评估，以找到适合于大规模数控系统的网络设计。最后，作者展示了异构连接对带宽的影响，以及不同的神经元映射算法如何增强这种影响。随着异构神经网络的使用，映射算法的需求也出现了，这也是模拟器中包含的一个功能。

#### **1104, 噪声对神经形态计算应用中 LIF 神经元模型的影响**

Thi Kim Thoa Thieu, Roderick Melnik, 2022.5.18

人工神经网络(ANN)已被广泛用于描述生物系统产生的问题和构建神经形态计算模型中。受生物神经元启发的第三代神经网络——脉冲神经网络(SNNs)能够更真实地模拟人脑。这些领域的一大类问题的特点是必须通过整合和激发神经元模型来处理神经元、棘突和突触的组合。受在生物医学研究神经形态计算中神经元

整合和激发的重要应用的启发，本文工作的主要重点是分析加性和乘性类型的随机输入电流以及随机不应期对 LIF 突触传导神经元模型的影响。本文作者提出并描述了具有随机输入的 LIF 突触电导模型。在数值环境中使用基于 Langevin 随机动力学的描述，作者分析了细胞膜电位中噪声的影响。具体而言，作者提供了模型的细节以及代表性的数值示例，并讨论了随机输入对细胞膜电位的时间演变、神经元的相应尖峰活动和放电率的影响。数值结果表明，在 LIF 突触传导模型中，随机输入强烈影响神经元的尖峰活动。此外，作者观察到乘法噪声的存在导致 LIF 突触传导动力学中的突发放电。然而，当增加泄漏电导值时，爆裂情绪会降低。当平均注入电流值足够大且不应期的标准偏差增加时，尖峰列的不规则性增加。随着更多不规则的尖峰序列，作者预期 LIF 突触传导系统中的突发减少。LIF 神经元中的随机输入可以减少神经元对 SNN 或 ANN 系统中的每个刺激响应。更好地理解神经网络系统中的不确定性因素可能有助于 SNN 算法的进一步发展，用于更高层次的脑启发功能研究和其他应用。

### 1105, 用于神经形态计算的 ZnSnO<sub>y</sub>/ZnSnO<sub>x</sub> 双层透明记忆突触装置

Dayanand Kumar、Aftab Saleem、Yeong Her Wang, 2022. 8. 8

IEEE

这篇论文主要研究了用于神经形态计算应用的透明双层忆阻器的突触特征。在人工突触应用中，忆阻器的应用潜力很大，主要因为其结构简单、纳秒级的优异开关速度、低功耗、长保留时间、出色的长周期耐久性和 CMOS 兼容性等特点。该忆阻突触装置是在 ITO 涂层玻璃板上制造的，在混合 Ar 和 O<sub>2</sub> 中通过 RF 溅射在 ITO 底部电极 (BE) 上沉积 15nm 厚的 ZnSnO<sub>x</sub> 氧化物层，沉积压力为 10mTorr，氧气

百分比为 3%。然后通过 RF 溅射也沉积了 5nm 厚的 ZnSnO<sub>y</sub> 层，其中沉积压力使用 20mTorr 而不是 10mTorr 并且氧百分比从 3%增加到 6%。最后在通过 RF 溅射沉积了 80nm 厚的 ITO 顶电极 (TE)。文中实验说明了该突触器件在 400 - 800nm 可见光波长范围内的平均透射率约为 85%，还说明在生物神经子系统中从突触前神经元到突触后神经元的信息传输由 Ca<sup>2+</sup>离子(递质)通过生物突触的激发机制控制，这样造成的后果就是突触被调节。本文还对具有逐渐置位和复位器件的双极 I-V 特性进行了研究，器件电压分别为 0.8 和-1.0，并且为了进一步取人存储器的性能，还通过在 SET 条件下使用 1V 的脉冲高度，高速动态脉冲诱导 RS 特性和-1.2V 脉冲宽度为 40ns 的 RESET 条件下测量了忆阻器件的可靠性特性。实验说明该装置在线性和重复 P 和 D 周期方面表现出优异的神形态特征，其中线性 P 和 D 为 1.2 和-1.4、重复稳定耐久性为 5000 次，该装置的突触特征证明了它非常适合神形态计算应用。

### 1106, 基于会话推荐的通用增强可解释框架与知识图谱

Huizi Wu、Hui Fang、Zhu Sun、Cong Geng、Xinyu Kong、Yew-Soon Ong, 2022. 12. 14

上海财经大学、新加坡高性能计算研究所和前沿人工智能研究中心、蚂蚁集团

基于会话的推荐(SR)近年来受到越来越多的关注。相当多的研究致力于设计复杂的算法来提高推荐性能，其中深度学习方法占了大部分。然而，这些方法大多是黑盒模式的，忽略了提供适度的解释来方便用户理解，从而可能导致用户满意度降低，系统收益减少。因此，在我们的研究中，我们提出了一个通用的增强可解释框架，用于基于会话的推荐(即 REKS)，该框架致力于用马尔可夫决策过程改进现有的黑盒 SR 模型(标记为不可解释模型)。特别地，我们构造了一个具有会话

行为的知识图，并将 SR 模型作为马尔可夫决策过程的策略网络的一部分。基于我们特别设计的状态向量、奖励策略和损失函数，基于强化学习 (RL) 的框架不仅提高了推荐精度，同时还提供了适当的解释。最后，我们在五个具有代表性的、最先进的 SR 模型 (即 GRU4REC, NARM, SR-gnn, GCSAN, BERT4REC) 中实例化了 REKS，从而在四个数据集上对这些方法进行了广泛的实验，证明了我们的框架在推荐和解释任务上的有效性。

### **1107, 使用增强型知识图谱转换器进行跨域方面术语提取**

Phillip Howard、Arden Ma、Vasudev Lal、Ana Paula Simoes、Daniel Korat、Oren Pereg、Moshe Wasserblat、Gadi Singer、2022.10.18

美国英特尔实验室、以色列英特尔实验室、

方面术语的提取是文本精细粒度情感分析的关键步骤，当训练核测试数据来自同一域时，这项任务现存的方法有着不俗的表现。但是，当测试数据于训练数据来自不同域时，这些方法的效果急剧下降。为了解决这些方法可扩展性和健壮性不足的问题，我们提出了一种新的构建特定域的包含方面术语标识相关的信息知识图谱的方法。我们引入了一种从这些知识图谱中抽取信息将其注入到转换模型中的方法，该方法包含两种可选的知识插入机制：通过查询扩充以及控制关注模式。我们使用我们的方法在跨域的方面术语提取中得到了极好的表现，另外我们还调查了存在对转换效果有影响外部知识总量。

### **1108, 用户多方协作强化学习的分级策略**

Majd Ibrahim Ammar Faya, 2022.12.14

充分制定代理人行为的策略对于解决合作 MARL 问题至关重要。在这个领域中，一种直观有益但不常见的方法是预测代理的未来行为并据此进行规划。利用这一点，我们提出了一种两级层次结构，将新的信息理论目标与轨迹预测模型相结合，以学习“策略”。为此，我们引入了一种潜在策略，该策略学习两种类型的潜在策略：个体  $z_A$  和关系  $z_R$ ，使用修改的图形注意力网络模块来提取交互特征。我们鼓励每个代理通过在  $z_A$  上调节其局部 Q 函数来根据策略行事，并且我们进一步为代理配备了在  $z_R$  上调节的共享 Q 函数。此外，我们引入了两个正则化器，以使预测的轨迹准确且有回报。对谷歌研究足球 (GRF) 和星际争霸 (SC) II 微观管理任务的经验结果表明，我们的方法建立了一个新的技术状态，据我们所知，这是第一个 MARL 算法，用于解决所有超硬 SC II 场景以及 GRF 全游戏，胜率高于 95%，因此优于所有现有方法。

### 1109, MA-GCL: 图对比学习的模型增强技巧

Xumeng Gong, Cheng Yang, Chuan Shi, 2022. 12. 15

对比学习 (Contrastive learning, CL)，可以提取不同对比视图之间的共享信息，已经成为视觉表示学习的一个流行范式。受计算机视觉成功的启发，最近的工作将对比学习引入到图的建模中，被称为图对比学习 (GCL)。然而，在图中生成对比性视图比在图像中生成对比性视图更具挑战性，因为我们对如何在不改变图的标签的情况下大幅增加图的内容几乎一无所知。我们认为，GCL 中典型的数据增强技术（如边缘丢弃）不能产生足够多的对比性视图来过滤掉噪音。此外，以前的 GCL 方法采用了两个具有完全相同的神经结构和绑定参数的视图编码器，这进一步损害了增强后的视图的多样性。为了解决这一局限性，我们提出了一种

名为模型增强 GCL (MA-GCL) 的新范式，它将重点放在操纵视图编码器的架构上，而不是扰乱图形输入。具体来说，我们为 GCL 提出了三种易于实现的模型增强技巧，即不对称、随机和洗牌，它们分别有助于缓解高频噪声、丰富训练实例和带来更安全的增强。这三个技巧都与典型的数据增强兼容。实验结果表明，通过在一个简单的基础模型上应用这三个技巧，MA-GCL 可以在节点分类基准上取得最先进的性能。

**1110, 可信医学人工智能的上下文相关可解释性和可争议性：早产儿发病率识别模型的错误分类识别**

Isil Guzey、Ozlem Ucar、Nukhet Aladag Ciftdemir、Betul Acuna, 2022. 12. 17  
特拉卡理工大学

尽管人工智能的机器学习 (ML) 模型在医学上取得了很高的性能，但它们并非没有错误。赋予临床医生识别不正确的模型建议的权力，对于建立对医学人工智能的信任至关重要。可解释的人工智能 (XAI) 旨在通过澄清人工智能的推理来支持终端用户来满足这一要求。最近有几项关于生物医学成像的研究取得了令人鼓舞的成果。尽管如此，使用表格数据的模型的解决方案还不足以满足临床医生的要求。本文提出了一种方法来支持临床医生识别用表格数据训练的 ML 模型的失败。作者把方法建立在三个主要的支柱上：利用临床背景的潜在空间来分解特征集，评估全局解释的临床关联，以及基于潜在空间相似性 (LSS) 的局部解释。作者在基于 ML 的感染引起的早产婴儿病症识别上展示了他们的方法。由于模型失败而导致的死亡、终身残疾和抗生素耐药性的风险是该领域的一个开放性研究问题。作者用他们的方法实现了识别两个模型的错误分类案例。通过对当地解释的语境

化，文章提出的解决方案为临床医生提供了可操作的见解，以支持他们自主地做出明智的最终决定。

### 1111, 基于熵的神经网络逻辑解释

Pietro Barbiero<sup>1</sup>、Gabriele Ciravegna、Francesco Giannini、Pietro Li、Marco Gori、Stefano Melacci<sup>3</sup>, 2022

剑桥大学、佛罗伦萨大学、锡耶纳大学、蔚蓝海岸大学

自从立法者开始要求安全关键领域的可解释模型以来，可解释的人工智能迅速出现。基于概念的神经网络已经作为可解释的设计方法出现，因为它们利用人类可理解的符号（即概念）来预测类成员资格。然而，这些方法中的大多数都侧重于识别最相关的概念，但并未对分类器如何利用这些概念进行预测提供简明、正式的解释。在本文中，我们提出了一种新颖的端到端可微分方法，能够使用一阶逻辑的形式从神经网络中提取逻辑解释。该方法依赖于基于熵的标准，该标准可自动识别最相关的概念。我们考虑了四个不同的案例研究来证明：(i) 这种基于熵的标准能够在从临床数据到计算机视觉的安全关键领域中提取简洁的逻辑解释；(ii) 所提出的方法在分类准确性方面优于最先进的白盒模型。

### 1112, 在放射学计算机辅助诊断的深度学习系统中使用可解释性的系统综述：可解释性 AI 的局限性？

Arjan M. Groen、Rik Kraan、Shahira F. Amirkhan、Joost G. Daams、Mario Maas, 2022. 12

阿姆斯特丹运动科学放射与核医学部、阿姆斯特丹医学图书馆、阿姆斯特丹运动

## 科学放射与核医学部

最近在放射学计算机辅助诊断(CAD)方面的研究越来越多地采用了深度学习(DL)。与其他形式的CAD相比,DL的优势在于它可以自动学习从高维图像中检测视觉特征[2]。因此,这些系统不依赖于从医学图像中提取信息性视觉特征(即纹理、图案和形状)的手工特征或无线电软件。与基于特征的方法不同,DL可以端到端工作,这意味着它可以在一端接收放射图像,处理图像,并在另一端进行诊断。由于这种方法将特征检测的任务完全委托给学习算法,端到端DL是一种用于从图像中提取和微调视觉特征的高度可扩展的方法。由于其参数数量众多,通常在数百万量级,DL模型通常被称为黑盒。在临床应用中,这导致了几个问题。例如,端到端DL模型可能会学习到有偏差的模式,并且无法对分布外数据执行。此外,缺乏解释性的DL系统可能不会被患者和护理人员视为值得信赖,并且可能无法遵守自动决策系统强制解释的规定。为了解决这些问题并帮助社会采用深度学习,一些研究人员提出了可解释AI(XAI)作为DL系统设计的组成部分。

### 1113, FaceTopoNet: 使用面部拓扑学习的面部表情识别

Mojtaba Kolahe Douzi, Alireza Sepas-Moghaddam, Ali Etemad, 2022. 9. 16

现有研究表明,使用顺序学习器学习面部不同组成部分的顺序可以在面部表情识别系统的性能中发挥重要作用。本文提出了FaceTopoNet,FaceTopoNet是一种用于面部表情识别的端到端深度模型,它能够学习面部的有效树形拓扑结构。模型遍历学习树以生成序列,然后使用该序列形成嵌入以提供序列学习器。本文对四个大规模的野外面部表情数据集(即AffectNet、FER2013、ExpW和RAF-DB)和一个实验室控制的数据集(CK+)进行了大量实验,以评估效果,实验结果表明

FaceTopoNet 在其中三个数据集上有着目前最先进的性能，并在其他两个数据集上获得了有竞争力的结果。本文也进行了严格的消融和敏感性实验，以评估模型中不同组件和参数的影响。本文最后进行了鲁棒性实验，并证明与该领域的其他领先方法相比，FaceTopoNet 对遮挡的鲁棒性更强。

#### 1114, AlphaFold 蛋白质结构数据库中预测后修改的案例

Haroldas Bagdonas、Carl A. Fogarty、Elisa Fadda、Jon Agirre, 2021.10.29

University of York、Maynooth University

AlphaFold2 的到来彻底改变了结构生物学的工作流程。然而，该算法没有考虑影响蛋白质结构和功能的基本修改，这给了我们部分图片。在这里，我们讨论了如何以相对直接的方式解决这种遗漏，从而对复杂的生物分子系统进行完整的结构预测。DeepMind 和 EMBL-EBI 最近发布的 AlphaFold 蛋白质结构数据库 1 标志着结构生物学的重大突破，因为它为全球科学界提供了对来自人类的 20,000 种蛋白质和来自 20 种其他生物学相关生物体的蛋白质的高度准确的结构预测。包括大肠杆菌。像许多研究大分子结构的科学家一样，我们对这一发展感到由衷的兴奋，但我们认为，以目前的形式误解其内容的可能性不可忽视。特别是，AlphaFold 数据库中的仅蛋白质预测意味着辅因子，最重要的是，由于技术的范围，翻译后修饰被排除在外是可以理解的。最相关的共翻译和翻译后修饰是蛋白质糖基化——相关且非常明显，正如最近对完全糖基化 SARS-CoV-2 刺突蛋白动力学的研究所说明的那样 2, 3。事实上，在这 20,000 种预测的人类蛋白质中，有 50% 到 70% 被认为是糖基化的 4，但这些都没有在数据库中明显突出显示。

**1115, 基于网络抓取的多通道数据的对比语言视觉人工智能模型显露出性客观化倾向**

Robert Wolfe, Yiwei Yang, Bill Howe, Aylin Caliskan, 2022. 12. 22

华盛顿大学

通过对比语言-图像预训练 (CLIP) 目标, 对九个在网页剪贴板上训练的语言-视觉人工智能模型进行评估, 以寻找心理学家研究的偏见证据: 对女孩和妇女的性物化, 当一个人的人性特征被忽视, 这个人被当作一个身体或身体部分的集合时发生。第一个实验使用了来自性客体化和情感数据库 (Sexual Objectification and Emotion Database) 的标准化女性形象, 结果发现, 与心理学之前的研究一致, 人类特征与客体化女性形象没有关联: 模型对情绪状态的识别是由对象是完全穿着还是部分穿着来调节的。嵌入式关联检验 (EAT) 对愤怒 ( $d > 0.8$ ) 和悲伤 ( $d > 0.5$ ) 都给出了显著的效应值。第二个实验测量代表性应用中的效果: 自动图像字幕器 (南极字幕) 包括表示情感的词, 对于部分穿着的妇女的图像, 其频率少于对于完全穿着的妇女的图像的 50%。第三个实验发现, 相对于男性专业人士的形象, 女性专业人士 (科学家、医生、行政人员) 的形象更容易与性描述联系在一起。第四个实验表明, 对于 VQGAN-CLIP (17 岁), “a [age] year old girl” 的提示产生了高达 73% 的时间的性感化图像 (如由 NSFW 分类器确定的), 对于 Stable Diffusion (14 和 18 岁), 产生了高达 40% 的时间的性感化图像; 男孩的相应比率从未超过 9%。有证据表明, 在自动收集的网页剪贴板上训练的语言视觉人工智能模型会学习性客观化的偏见, 并传播到下游应用程序。

### 1116, THREATKG: 用于自动化开源网络威胁情报收集和管理的威胁知识图谱

Peng Gao, Xiaoyuan Liu, Edward Choi, Sibon Ma, Xinyu Yang, Zhengjie Ji, Zilin Zhang, Dawn Song, 2022. 12. 20

加利福尼亚大学伯克利分校

尽管越来越多地采用开源网络威胁情报(OSCTI)来获取有关网络威胁的知识,但很少有人努力从大量可用的非结构化 OSCTI 报告(例如,安全文章,威胁报告)中获取知识。这些报告提供了各种实体(例如 ioc、威胁行为者、ttp)和关系的全面威胁知识,但由于报告格式多样、报告数量大、自然语言报告文本结构复杂和细微差别,这些知识很难收集。为了弥合差距,我们提出了一个用于自动化开源网络威胁知识收集和管理的系统 THREATKG。THREATKG 自动收集各种来源的大量 OSCTI 报告,提取高保真的威胁知识,构建威胁知识图,并通过不断吸收新知识来更新知识图。为了应对多重挑战,THREATKG 提供了:(1)用于建模各种威胁知识实体和关系的层次本体;(2)基于深度学习的精确威胁知识提取管道;(3)用于威胁知识图构建、持久化、更新和探索的可扩展系统架构。通过对大量报告的评估,证明了 THREATKG 在威胁知识收集和管理方面的有效性。

### 1117, 预测原子候选 OWL 类公理的得分

Ali Ballout, Andrea G. B. Tettamanzi, C'elia da Costa Pereira, 2022. 12. 21

抽象候选公理评分是根据已知事实或数据提供的证据评估候选公理的可接受性的任务。对候选公理进行可靠评分的能力是自动模式或本体归纳所必需的,但它对本体和/或知识图验证也很有价值。准确的公理评分启发式算法通常在计算上很昂贵,如果您希望在迭代搜索技术中使用它们,例如逐级生成和测试或进化算

法，这需要对大量候选公理进行评分，这是一个问题。我们解决了开发一个预测模型作为推理的替代品的问题，该模型预测候选类公理的可能性分数，并且足够快，可以在这种情况下使用。为此，我们使用从本体的包容结构中获取的语义相似性度量。我们表明，本工作中提供的方法可以准确地学习候选 OWL 类公理的可能性分数，并且可以对各种 OWL 类的公理进行学习。

### 1118, 利用知识库的先验知识对结构数据源进行自动语义建模

Jiakang Xu, Wolfgang Mayer, HongYu Zhang, Keqing He, Zaiwen Feng

在线分享语义内容的一个关键步骤是将结构性数据源映射到公共领域的本体论。这个问题被表示为关系-本体映射问题 (Rel2Onto)。手动建立数据的语义模型需要巨大的努力和专业知识。因此，一个自动学习数据源语义的方法是可取的。大多数现有的工作都是研究源属性的语义注释。然而，虽然很关键，但自动推断属性之间关系的研究却非常有限。在本文中，我们提出了一种新的方法，用于对结构化数据源进行语义注释，使用机器学习、图匹配和修正的频繁子图挖掘来修正候选模型。在我们的工作中，知识图被用作先验知识。我们的评估表明，在只有少数语义模型已知的棘手情况下，我们的方法优于两个最先进的解决方案。

### 1119, 双相位共轭反馈半导体纳米激光增强神经形态储层计算系统的预测性能

Xing Xing Guo, Shui Ying Xiang, YAi Jun Wen, Yue Hao 等, 2021.12.22

首次提出了一种基于双相位共轭反馈(PCF)半导体纳米激光器(SNL)的神经形态储层计算(RC)系统，并进行了数值验证。通过 Santa Fe 混沌时间序列预测任务研究了该 RC 系统的预测性能。在速率方程中引入了 Purcell 腔增强的自发发射

因子  $F$  和自发发射耦合因子  $\beta$ ，并广泛分析了  $F$  和  $\beta$  对 RC 系统预测性能的影响。为了进行比较，还考虑了单 PCF 下基于 snl 的 RC 系统的预测性能。仿真结果表明，与单 PCF 的 SNL-based RC 系统相比，双 PCF 的 SNL-based RC 系统具有更好的预测性能。此外，还考虑了偏置电流、输入信号调制深度、反馈强度以及反馈延迟的影响。本文提出的基于 snl 的双 PCF 的 RC 系统具有开发基于 RC 的神经形态光子集成电路的潜力。

本文首次提出了一种基于双相位共轭反馈的半导体纳米激光器的神经形态储层计算方法，并对其进行了数值验证。仿真结果表明，与单 PCF 的 snl RC 系统相比，具有双 PCF 的 snl RC 系统具有更好的预测性能。这可能是因为第二个反馈引入了更丰富的动态。此外，本文还提出了偏置电流、输入信号调制深度、反馈强度和反馈延迟是获得较好 RC 预测性能的关键因素。此外，我们还发现，较大的 Purcell 空腔的自发辐射因子  $f$  增强，且较大的自发辐射因子  $f$  增强发射耦合因子  $\beta$  可提高预测性能，因为较高的发射耦合因子  $f$  和  $\beta$  可增强松弛振荡阻尼。因此，本文提出的基于 snl 的双 PCF RC 可为基于 RC 的集成神经形态光子系统的设计提供理论指导。

## 1120, 神经形态计算的设计技术协同优化

Ankita Paul, Shihao Song, Anup Das, 2021. 12. 15

本文作者提出了在基于非易失性存储器 (NVM) 的多核神经形态硬件的处理核心上实现机器学习推理的设计技术权衡分析，以研究采用 NVM 作为突触存储的神经形态硬件的推断寿命。分析的一个基本观察结果是，NVM 细胞的读取耐久性取决于其编程的突触重量和暴露于细胞的电压。通过对缩放工艺技术节点的详细电路

级仿真,本文显示了缩放设计对 NVM 的读取耐久性的负面影响,这直接影响了 NVM 的推断寿命。在更精细的粒度上,核心的推断寿命取决于在核心上编程的突触权重的电阻状态和由电流路径上的寄生组件引入的核心内部的电压变化。作者的分析表明,由于电流路径上的寄生分量,暴露于交叉开关内 NVM 单元的电压会发生变化,这导致交叉开关中 NVM 单元之间的非对称读取耐久性。作者表明,这样的设计和技术特征可以结合到设计流程中,从而显著提高推断寿命。

### 1121, 终身可靠性感知神经形态计算案例

Shihao Song 、Anup Das, 2022. 12. 20

美国宾夕法尼亚州费城德雷塞尔大学计算机工程

在非易失性存储器上使用神经形态计算能够让使用脉冲计算和仿生学习算法的机器学习任务的性能提高以及能耗降低。但是操作非易失性存储器的高电压会加速神经元 CMOS 电路的老化。本文评估了 10 种前沿的机器学习应用对于可靠性和性能之间权衡的影响,并在其中考虑了负偏置温度不稳定性(NBTI)和时变击穿(TDDB)两种故障机制。本文为了提高神经形态电路的终身可靠性,引入了周期性松弛神经元的方法,即周期性的对神经元进行减压。本文主要的性能衡量指标为脉冲间隔周期(inter-spike interval, ISI)和紊乱脉冲个数(disorder spike count, DSC),而与神经元周期性松弛的指标主要为减压间隔(tDSI)和减压完成时间(tDSC)。本文的实验主要是将 tDSI 从 10ms 增加到 50ms 并观察各指标的变化。对于可靠性,实验结果有三点: 1) NBTI 和 TDDB 的老化会随着 tDSI 的增加而增加。2) tDSI 对于老化的影响有着应用差异性,即对于不同的机器学习应用 tDSI 会有着不同程度的影响。3) 相比于 NBTI, TDDB 对于不同应用的老化增加程

度很接近。对于性能，实验结果表示：随着 tDSI 的增加 ISI 失真和 DSC 都会减少。对于热影响，实验结果表示在 325K 和 350K 的温度下的老化会比在 300K 温度下的老化增加 7%和 26%。

### 1122, 学习解释：快速生成稳定的解释

Xuelin Situ<sup>1</sup>、Ingrid Zukerman<sup>1</sup>、Cecile Paris、Sameen Maruf、Gholamreza Haffari<sup>1</sup>, 2021

莫纳什大学、澳大利亚联邦科学与工业研究组织创新部门

解释机器学习模型，特别是黑盒模型的结果的重要性已被广泛认可。最近的方法是通过识别输入特征对该结果的贡献来解释一个结果。在涉及大型黑盒模型或复杂输入的环境中，这导致了对计算的要求。此外，这些算法经常受到低稳定性的影响，其解释在类似的例子中变化很大。在本文中，我们提出了一种学习解释（Learning to Explain, L2E）的方法，从所有的训练实例中同时学习基础解释算法的行为。一旦解释算法被提炼成解释者网络，它就可以被用来解释新的实例。我们在三个分类任务上的实验，将我们的方法与六种解释算法进行了比较，结果表明，L2E 比这些算法快 5 到  $7.5 \times 10^4$  倍，同时产生了更稳定的解释，并且对黑盒模型有相当的可信度。

### 1123, 基于 CNN 和可解释人工智能的孟加拉手写数字识别

Md Tanvir Rouf Shawon、Raihan Tanvir、Md. Golam Rabiul Alam, 2022. 12. 23  
Ahsanullah 科技大学，布拉克大学

手写字符识别是当今研究的一个热门话题。如果我们能用光学字符识别（OCR）技

术将一张手写的纸转换成可搜索的文本文件，我们就可以很容易地理解内容，而不需要阅读手写的文件。英语中的 OCR 非常普遍，但在孟加拉语中，很难找到一个高质量的 OCR 应用。如果能将机器学习和深度学习与 OCR 合并，这可能是对这个领域的巨大贡献。不同的研究人员已经提出了许多识别孟加拉语手写字符的策略。在他们的工作中使用了大量的 ML 算法和深度神经网络，但他们的模型并不存在解释。在本文的工作中，作者使用了各种机器学习算法和 CNN 来识别手写的孟加拉语数字。作者从一些 ML 模型中得到了可接受的准确性，而 CNN 则提供了很大的测试准确性。Grad-CAM 被用作本文中的 CNN 模型的 XAI 方法，这让我们对该模型有了深入的了解，并帮助我们检测出从图像中识别数字的关注点。

#### **1124, XAI 手册：迈向可解释 AI 的统一框架**

Sebastian Palacio、Adriano Lucieri、Mohsin Munir、Jörn Hees、Sheraz Ahmed、Andreas Dengel, 2021. 5. 14

德国人工智能研究中心，德国凯泽斯劳滕

可解释人工智能 (XAI) 领域已迅速成为一个繁荣和多产的社区。然而，这一领域一个沉默、反复出现和公认的问题是对其术语缺乏共识。特别是，每一项新的贡献似乎都依赖于自己（通常是直觉的）版本的术语，如“解释”和“解释”。这种混乱阻碍了该领域在实现科学和监管要求方面的进展的整合，例如，在比较方法或确定其是否符合 r. t. 偏见和公平约束时。本文提出了一个理论框架，不仅为这些术语提供了具体的定义，而且还概述了产生解释和解释所需的所有步骤。该框架还允许对现有贡献进行重新组织，以便能够衡量其范围，从而使其与其他方法相比较。本文表明，该框架在解释、可解释性和评估指标方面符合需求。本文

提供了一个用例，展示了该框架如何用于比较 LIME、SHAP 和 MDNet，并确定了它们的优势和缺点。最后，本文从框架的角度讨论 XAI 的相关趋势以及未来工作的建议。

## 1125, 用于峰值神经拟态计算的被动记忆交叉条的有效信号传递

Ali Shiri Sichani、Kishore Kumar Kadari、Wilfrido A. Moreno, 2021.12.29  
University of Memphis、University of South Florida

各种各样的材料系统，比如新兴的非易失性存储器 (envm)，可以通过信号变化来有效地模拟突触行为。这些材料系统通常表现出记忆系统的特征，包括电阻性随机存取存储器 (RRAM)、相变材料 (PCM)、自旋电子和铁电器件等。需要明确的信号策略和范例来使用被动记忆 RRAM 设备来提高神经形态计算系统，这是专注于神经拟态计算系统开发有效的信号的重要原因。

因为可以实现高密度的组装，被动记忆交叉棒成了众多神经拟态处理器中的一个很好的选择。PCM 器件的电阻调制可以通过调节硫系材料中结晶体积与非晶体积的比值实现。此外，PCM 上的电压调制可以改变结晶和非晶化比，从而使得 PCM 器件具有模拟生物学上范畴中的合理学习的潜力。本文所提出的 PCM 神经拟态系统具有 2 个 PCM，其结构为 1 个晶体管配上 1 个忆阻器 (1T-1M) 增强后神经元，以及另一个 1T-1M 抑制后神经元。这种结构实现了记忆交叉条的密集实现，但是同时也增加了制造的复杂性。无源记忆交叉条的实现在峰值计算和传统神经网络平台上具有显著优势，同时也无需在忆阻器上实现额外的设备，即 0T-1R 和数字外围电路来克服电路挑战。记忆 RRAM 器件在模拟生物突触和超高密度集成制造方面表现出良好的模拟功能。由于记忆性 RRAM 作为突触元件的优异性能，Spike

Timing Dependent Plasticity (STDP) 学习的信号传导方法一直关注于记忆性 RRAM 范式。要找到一种为峰值神经形态计算提供比例信号的通用方法，需要从各个方面评估记忆装置，其解释如下：1-信号幅度对长期抑郁 (LTD) 和长期增强 (LTP) 的影响，2-频率对 memductance 调制的影响及其演化步骤，3-在记忆交叉条上寻找自适应现象，以证实被动记忆装置对 spike 神经形态计算的适用性。这些要求促使作者在本论中提出了这种评估和方法。作者设计并展开了相关的实验。结果表明，可以设计一种有效的训练信号，使记忆装置在某些条件下用于神经拟态计算。比例信号为被动记忆交叉条提供了学习，并为节能计算铺平了道路。此外，记忆电阻器在脑启发计算中更重要的特性是它们的自适应能力。

## 1126, 用于神经形态计算的生物 SFQ 电路家族的最新进展

Vasili K. Semenov, Evan B. Golden, Sergey K. Tolpygo, 2022.12.27

超导体单通量量子 (SFQ) 技术由于低能量耗散和高时钟频率 (可能高达 100GHz) 的特点，对神经形态计算具有巨大前景。本文作者最近提出了一个新的生物 SFQ 电路家族，其中信息存储为超导回路中的电流值，并以在回路之间传播的 SFQ 脉冲的速率传输。在处理正数的最简单情况下，这种方法需要单线传输通道。在双极数字的更一般情况下，它需要双轨传输通道。为此，本文作者开发了一种新型的双轨输出比较器。该比较器是双极乘法器的重要组成部分，该乘法器也经过了设计、制造和测试。作者演示了所提出的生物 SFQ 电路和本工作中演示的模拟双极乘法器的神经形态电路示例。结果表明，使用所提出的电路，非常接近于实现广泛的模拟功能。可以通过重新路由 SFQ 脉冲的反馈流来重新编程期望的功能。本文中，作者讨论了其提出的生物 SFQ 方法的战略优势，例如，生物 SFQ 细胞的

固有异步特性，不需要明确的时钟信号。因此，生物 SFQ 电路没有竞速误差，并且能够容忍传播 SFQ 脉冲的偶然碰撞。这种容差是由于比较器在其灰色区域内操作所产生的数据信号的随机性质。这些电路是在麻省理工林肯实验室为超导体电子学开发的八钨层制造工艺 SFQ5ee 中制造的。

## 1127, 神经形态系统用非晶态金属氧化物半导体薄膜的多层交叉栅阵列

Etsuko Iwagi Takumi Tsuno、Takahito Imai、Yasuhiko Nakashima、Mutsumi Kimura, 2022. 8. 31

IEEE

目前，神经形态系统的应用开发中使用了非晶金属氧化物半导体 (AOS) 薄膜开发了多层交叉阵列，这种阵列可以实现多层结构，主要原因是因为 AOS 薄膜可以通过简单的溅射方法沉积，而且并不需要热处理，这样就不会损坏下层结构。主要过程如下：首先，通过蒸镀法沉积 Au 薄膜并将其作为电极，之后通过溅射法沉积非晶 In-Ga-Zn-O ( $\alpha$ -IGZO) 薄膜作为电导变化层，重复这些过程，并完成多层交叉阵列，其中三个电导变化层中的每一个都夹在电极之间。接下来，将多层交叉开关阵列实施到具有改进的 Hebbian 学习的神经形态系统中，这能够让系统在没有控制电路的情况下进行自主学习，并且确认了关联记忆功能，并且还保证了更高级功能实现的可能性。本文实验中为确认联想记忆功能在训练阶段对字母“T”和“L”进行了学习，实验结果显示“T”和“L”的字母字符已成功学习，即“T”与“L”字母字符已被成功复制，但是有一个故障示例，这可能是由于 AOS 薄膜多层交叉阵列的模拟忆阻特性导致。结果表明在任何情况下，都可以说是确认了联想记忆功能，但值得注意的是虽然电阻变化缓慢，但联想记忆功能在实际时间内

得到了确认。这些结果导致未来神经形态系统中突触元件的天文大规模集成 (LSI)。

### 1128, 面向大规模零次学习的语义增强知识图谱

Jiwei Wei、Yang Yang、Zeyu Ma、Jingjing Li、Xing Xu、Heng Tao Shen,

2022.12.26

零镜头学习一直是视觉和语言领域的研究热点。目前已有的方法大多采用结构化知识信息对类别之间的显式关联进行建模,利用深度图卷积网络在不同类别之间传播信息。但是,在已有的结构化知识图中很难添加新的类别,深度图卷积网络存在过平滑问题。本文提出了一种新的语义增强知识图谱,它既包含专家知识,又包含类别语义关联。语义增强的知识图谱可以进一步增强类别之间的相关性,便于吸收新的类别。为了在知识图上传播信息,提出了一种新的残差图卷积网络 (Residual graph Convolutional Network, ResGCN),有效地缓解了知识图的过平滑问题。在广泛使用的大规模 ImageNet-21K 数据集和 AWA2 数据集上的实验证明了该方法的有效性,建立了一种新的零镜头学习技术。此外,我们在具有各种特征提取网络的大规模 ImageNet-21K 上的结果表明,我们的方法具有更好的泛化和鲁棒性。

### 1129, 图形神经网络的对抗边缘检测

Xiaojun Xu Yue Yu Hanzhang Wang Alok Lal Carl A. Gunter Bo

Li, 2022.12.27

图形神经网络 (GNNs) 已广泛应用于生物信息学、药物设计和社交网络等不同任

务。然而，最近的研究表明，GNN 容易受到对抗性攻击的影响，这些攻击旨在通过添加细微的扰动来误导节点（或子图）分类预测。特别是，通过添加/删除少量边缘，已经提出了针对 GNN 的几次攻击，这引起了严重的安全问题。由于扰动的小幅度和图形数据的离散性，检测这些攻击具有挑战性。在本文中，我们提出了一种通用的对抗性边缘检测流水线 EDoG，而不需要基于图生成的攻击策略知识。具体而言，我们提出了一种结合链接预测的新的图生成方法，以检测可疑的对抗边缘。为了有效地训练图生成模型，我们从给定的图数据中采样了几个子图。我们表明，由于对抗边的数量在实践中通常很低，因此采样的子图将包含基于并集的对抗边的概率很低。此外，考虑到干扰大量边缘的强攻击，我们提出了一组新的特征来执行离群点检测，作为检测的预处理。在三个真实世界的图形数据集上的广泛实验结果表明，EDoG 可以在不需要任何攻击类型知识的情况下，针对四种最先进的隐形攻击策略实现 0.8 AUC 以上的 AUC（例如目标受害者节点的程度）；并且在知道攻击类型的情况下约为 0.85。EDoG 显著优于传统的恶意边缘检测基线。我们还表明，充分了解我们的检测管道的自适应攻击很难绕过它。我们的结果揭示了提高 GNN 鲁棒性的几个原则。

### 1130, 结合符号和亚符号方法的网络攻击检测、解释和过滤

Anna Himmelhuber, Dominik Dold, Stephan Grimm, Sonja Zillner, Thomas Runkler, 2022.12.29

在网络安全领域的入侵检测方面，图结构数据的机器学习（ML）最近受到了更大的关注。由于监测工具产生的数据量越来越大，以及越来越复杂的攻击，这些 ML 方法正在获得关注。知识图谱及其相应的学习技术，如图谱神经网络（GNN），具

有使用人类可理解的词汇表无缝整合多个领域数据的能力，正在网络安全领域找到应用。然而，与其他连接主义模型类似，GNNs 在决策方面缺乏透明度。这一点尤其重要，因为在网络安全领域往往有大量的假阳性警报，这样的分流需要由领域专家完成，需要大量的人力。因此，我们正在解决 GNN 的可解释人工智能(XAI)问题，通过探索结合网络安全领域的符号和亚符号方法来加强信任管理，其中包含了领域知识。我们通过在一个工业演示系统中生成解释来实验这种方法。所提出的方法被证明可以为不同场景的警报产生直观的解释。解释不仅为警报提供了更深入的见解，而且还使假阳性警报减少了 66%，当包括保真度指标时，减少了 93%。

**1131, SoloGAN: 多域多模态非配对图像到图像翻译通过单一生成对抗网络**

Shihua Huang, Cheng He, Ran Cheng, 2022. 6. 30

尽管使用生成对抗网络(GAN)在图像到图像(I2I)转换方面取得了重大进展，但使用一对生成器和鉴别器将图像有效地转换为多个目标域中的一组不同图像仍然具有挑战性。现有的多模式 I2I 翻译方法针对不同领域采用多个特定领域的内容编码器，其中每个特定领域的内容编码器仅使用来自同一领域的图像进行训练。本文认为内容领域不变性的特征应该从所有领域的图像中学习，因此现有方案的每个特定领域的内容编码器都无法有效地提取领域不变的特征。为了解决这个问题，本文提出了一种灵活且通用的 SoloGAN 模型，用于在具有未配对数据的多个域之间进行高效的多模式 I2I 转换。与现有方法相比，SoloGAN 算法使用带有附加辅助分类器的单个投影鉴别器，并为所有域共享编码器和生成器，因此可以使用来自所有域的图像有效地训练 SoloGAN 模型，从而可以有效地提取域不变性内

容表示。针对 SoloGAN 模型的几个对应物和变体在广泛数据集上的定性和定量结果证明了该方法的优点，特别是对于具有挑战性的 I2I 翻译任务，即涉及极端形状变化或需要在之后保持复杂背景不变的任务翻译。在猫狗翻译任务的用户研究中，93.65%的用户更喜欢我们的合成图像，明显优于同类方法。

### **1132, PDBsum 附加功能: SARS-CoV-2 和 AlphaFold 模型**

Roman A. Laskowski、Janet M. Thornton, 2022. 9. 14

European Molecular Biology Laboratory

PDBsum 网络服务器提供蛋白质数据库 (PDB) 中条目的结构分析。此处描述了两个最近添加的内容。首先是对 PDB 中 SARS-CoV-2 病毒蛋白结构的详细分析。这些包括关注的变体，它们显示在蛋白质的序列和 3D 结构上。第二个添加是包含可用的人类蛋白质 AlphaFold 模型。这些页面允许通过结构注释序列 (SAS) 服务器根据 PDB 中的现有结构搜索蛋白质，因此可以轻松地将预测模型与实验确定的结构进行比较。

### **1133, 探索视觉 Transformer 作为扩散学习者**

He Cao<sup>1</sup>、Jianan Wang<sup>1</sup>、Tianhe Ren<sup>1</sup>、Xianbiao Qi、Yihao Chen<sup>1</sup>、Yuan Yao、Lei Zhang, 2022. 12. 29

粤港澳大湾区数字经济研究院、香港科技大学

基于分数的扩散模型已经引起了广泛的关注，并为最近的视觉生成任务的快速发展提供了资金。本文重点研究了以前被忽视的扩散模型骨架。我们系统地探讨了 Vision Transformers 作为扩散学习者的各种生成任务。通过我们的改进，基于

vanilla ViT 的骨干网 (IU-ViT) 的性能被提升到与传统的基于 U-Net 的方法相当。我们进一步提供了一个假设, 关于解开生成性主干作为编码器-解码器结构的含义, 并展示了概念验证实验, 验证了使用 ASymetric 编码器解码器 (ASCEND) 进行生成性任务的更强编码器的有效性。我们的改进在 CIFAR-10、CelebA、LSUN、CUB Bird 和大分辨率文本到图像任务上取得了有竞争力的结果。据我们所知, 我们是第一个在超过 64 x64 分辨率的文本到图像任务上成功训练单扩散模型的公司。我们希望这将激励人们重新思考基于扩散的生成式模型的建模选择和训练管道。

**1134, 双重问题:** 如何不使用由掩码语言模型合成的反事实来解释文本分类器的决定?

Thang M. Pham<sup>1</sup>、Trung Bui、Long Mai、Anh Nguyen, 2022

奥本大学、Adobe 研究院 (Adobe Research)

几十种归因方法背后的一个原则是, 将一个输入特征 (token) 被移除前后的预测差异作为其归因。一个流行的输入边际化 (Input Marginalization, IM) 方法 (Kim et al., 2020) 使用 BERT 来取代一个标记, 产生更可信的反事实。虽然 Kim 等人报告说 IM 是有效的, 但我们发现这个结论并不令人信服, 因为他们的论文中使用的 DeletionBERT 度量是偏向 IM 的。重要的是, 这种偏见存在于基于删除的衡量标准中, 包括插入、充分性和全面性。此外, 我们使用 6 个指标和 3 个数据集进行的严格评估发现, 没有证据表明 IM 比 "忽略不计" (L00) 的基线更好。我们发现 IM 不比 L00 好的原因有两个: (1) 从输入中删除一个词只能略微降低分类器的准确性; (2) 一个高度可预测的词总是被赋予接近零的属性, 不管它

对分类器的真正重要性。相比之下，通过 BERT 使 LIME 样本更加自然，在几个 ROAR 指标下持续提高 LIME 的准确性。

### 1135, 自然言论证中逻辑谬误的稳健和可解释识别

Zhivar Sourati、Darshan Deshpande、Himanshu Rawlani、Filip Ilievski、Hông-Ân Sandlin、Alain Mermoud, 2022.12.12

南加利福尼亚大学、瑞士国防装备采购局科学和技术部

在互联网时代，错误信息、宣传和有缺陷的论证的传播被放大了。鉴于数据量和识别违反论证规范的微妙性，用可以识别逻辑谬误的可靠方法来支持信息分析任务，如内容审核，是非常重要的。在文章中，作者将先前关于逻辑谬误的理论工作正规化为一个全面的三阶段评估框架，包括检测、粗粒度和细粒度分类。作者为每个阶段的评价调整现有的评价数据集；在原型推理、基于实例的推理和知识注入的基础上，设计了三个系列的稳健和可解释的方法。这些方法旨在将语言模型与背景知识和可解释机制相结合。此外，作者通过数据增强和课程学习的策略来解决数据稀少的问题。文章的三阶段框架原生地整合了先前的数据集和现有任务的方法，如宣传检测，作为一个总体的评估测试平台。作者在他们收集的数据集上广泛地评估这些方法，重点是它们的稳健性和可解释性。文章的结果提供了对这些方法在不同组件和谬误类别上的优势和劣势的洞察力，表明谬误识别是一项具有挑战性的任务，可能需要专门的推理形式来捕捉各种类别。作者在 GitHub 上分享了开源代码和数据，以支持逻辑谬误识别方面的进一步工作。

### 1136, DeepFace EMD: 使用逐块移土距离重新排序改进了分布外的人脸识别

Hai Phan、Anh Nguyen, 2022. 3. 25

奥本大学、卡耐基梅隆大学

人脸识别 (FI) 无处不在, 并推动执法部门做出许多重大决策。一种常见的 FI 方法通过获取图像嵌入之间的余弦相似性来比较两个图像。然而, 这种方法对未包括在训练集或库中的新类型的图像 (例如, 当查询脸被掩蔽、裁剪或旋转时) 的分布外 (OOD) 泛化较差。在这里, 本文提出了一种重新排序的方法, 该方法使用地球移动器的距离对图像块的深层空间特征进行比较。本文额外的比较阶段明确地在细粒度级别 (例如, 眼睛对眼睛) 检查图像相似性, 并且比传统的 FI 对 OOD 扰动和遮挡更加鲁棒。有趣的是, 在没有微调特征提取器的情况下, 本文的方法始终提高了所有测试的 OOD 查询的准确性: 掩蔽、裁剪、旋转, 同时在分布图像上获得类似的结果。

### 1137, 面向“通用用途”的脑启发计算系统

Youhui Zhang、Peng Qu, and Weimin Zheng, 2023. 1. 3

脑启发计算是指主要受大脑处理方式或结构启发的计算模型、方法和系统。最近的一项研究提出了“神经形态完备性”的概念和相应的系统层次结构, 这有助于确定脑启发计算系统的能力边界, 以及判断脑计算所涉及的软硬件是否相互兼容。神经形态计算 (neuromorphic computing) 一词首次出现在上世纪 90 年代由加州理工学院计算机科学教授卡弗·米德 (Carver Mead) 撰写的参考文献中。当时的神经形态计算系统是指利用非常大规模的集成电路, 用电子器件模拟生物神经系统的自适应大规模并行计算系统。它通常用基本物理现象作为计算原语, 用模拟

信号的相对值来表示信息。

随着时间的推移，神经拟态计算芯片的设计风格和使用技术发生了很大的变化，但它们的技术路线在一定程度上可以追溯到 Mead 教授的工作，主要有以下三个方面：

1. 生物大脑结构启发
2. 通过微/纳米器件模拟神经突触
3. 实现新型的低能耗且高效的智能电脑系统

作为一篇概述类论文，本文分析了现有的脑启发芯片设计特点和目前所谓的“通用”脑启发式计算应用开发框架，并介绍了该提案的背景和潜力。此外，通过与图灵完备性和近似计算的比较，以及与“通用”脑启发计算系统(这意味着计算系统可以支持所有可计算应用)的关系分析，介绍了这一概念的一些关键特征。最后，介绍了一种有前景的实现此类计算系统的技术途径，以及正在进行的研究和工作基础。作者认为这项工作有利于可扩展的神经形态完整硬件原语和相应芯片的设计。在此基础上，逐步实现“通用”脑计算系统，兼顾功能完备性和应用效率。

**1138**, 传感器和神经形态计算是实现高效节能计算机视觉所需的一切

Gourav Datta, Joe Mathai, Zihan Yin, Ajey P. Jacob, Akhilesh R. Jaiswal, Peter A. Beerel, 2022. 12. 21

由于高激活稀疏性，以及使用的累积 (AC) 而不是昂贵的乘法和累积 (MAC)，神经形态脉冲神经网络 (SNN) 已成为几种计算机视觉 (CV) 应用中传统 DNN 的低功耗替代方案。然而，大多数现有的 SNN 需要多个时间步骤来获得可接受的推理精

度，这阻碍了实时部署，增加了峰值活动，从而增加了能耗。本文作者提出了直接编码，其直接馈送 SNN 的第一层中的模拟像素值，以显著减少时间步长的数量。尽管具有直接编码的第一层 MAC 的开销对于深度 SNN 来说是可以忽略的，并且 CV 处理使用 SNN 是有效的，但是图像传感器和下游处理之间的数据传输消耗了大量带宽，并且可能增加总消耗。为了减轻这种担忧，本文作者提出了一种传感器内计算硬件和软件联合设计框架，用于 SNN 的图像识别任务。与传统 CV 处理相比，我们的方法将传感和处理之间的带宽减少了 12-96 倍，产生的总能量减少了 2.32 倍，ImageNet 上的准确度降低了 3.8%。本文作者的硬件修改可以很容易地集成到铸造厂可制造的 CMOS 图像传感器平台中，该硬件启发的算法修改显示在复杂视觉任务中的性能下降可以忽略不计。这项工作中报告的能量效益是基于前端传感器和视觉处理芯片紧密位于同一印刷电路板上的。一般来说，它们可以相隔很长的距离，需要长时间、高能耗的有线或无线数据传输（这在传感器融合和群体智能应用中很常见）。在这种情况下，整体能量改善将接近减少带宽的情况。

**1139**, 基于神经形态计算的铁电 FET 对称线性突触行为的设计技术协同优化

Guoqing Zhao、Shuhao Wu、Xuepeng Zhan、Mingfeng Tang 等, 2022.8.31

IEEE

目前，由于基于  $\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$  (HZO) 的铁电场效应晶体管 (FeFET) 具有很强的兼容性和可扩展性，因此在人工神经元方面得到广泛应用。本文通过对畴切换 (DS) 与电荷捕获之间的相关性 (CT) 的研究，并通过调制沟道电导来精确控制 FeFET 中的多态。通过该研究发现当

DS 贡献占主导时，在中间存储状态中可以很好地抑制周期间的变化，通过包括 CT

和温度影响进一步协同优化并实现了温度影响、对称线性电导调制和大电导比。本文为了模拟突触特征，采用了两种不同的脉冲方案，并通过同时施加脉冲  $V_g$  和  $V_d$  (50ns) 来实现对沟道电导的逐渐调制。实验结果中 CT 模式对应 CW 滞后回路，表明正向 IV 和反向 IV 曲线之间的正积分电荷，并且模式 A 中获得的 MW 大于模式 B 中的 MW。从实验结果还可以得出模式 A 中在不同温度下获得的四种不同状态的  $V_{th}$  变化。从  $V_{th}$  的统计信息的分析中，与 CT 模式相比，DS 模式的  $V_{th}$  变化较小。随着编程电压的逐渐增加，在读取操作之前，可以将更多的极化状态从向上切换到向下，而当温度从 350 K 下降到 245 K 时，观察到增强的存储能力。本文结果说明在 DS 和 CT 共存的情况下，HZO FeFET 实现了超过 3 位的稳定存储能力。当在脉冲方案、操作条件和测试温度中使用 DTCO，可以实现具有大电导比的对称线性突触特征，这证明了高可靠性 FeFET 在神经形态计算中的巨大潜力。

## 1140, 结构知识与常识知识的少镜头知识图谱补全研究

HAODI MA , 2022. 1. 3

佛罗里达大学

知识图谱 (KG) 已经成为各种自然语言处理应用的关键组成部分。常识知识图谱 (Commonsense knowledge graph, CKG) 是知识图谱的一种特殊类型，其中实体和关系由自由形式的文本组成。但是，以往的 KG 补全和 CKG 补全工作存在长尾关系和新增关系，这些关系没有太多的已知三元组可供训练。基于此，提出了利用图表示学习和少镜头学习优势的少镜头 KG 补全 (FKGC) 来挑战标注数据有限的问题。在本文中，我们以一系列方法和应用的形式，全面综述了前人在这方面的尝试。具体来说，我们首先介绍 FKGC 挑战、常用的 KG 和 CKG。然后系统地对现有

文献进行分类和总结,从文献类型和方法两方面进行分析。最后,我们介绍了 FKGC 模型在不同领域的预测任务中的应用,并对未来 FKGC 的研究方向进行了展望。

#### 1141, 基于句子变换器的多方面可解释归纳关系预测

Zhixiang Su、Di Wang、Chunyan Miao、Lizhen Cui, 2023.01.04

抽象候选公理评分是根据已知事实或数据提供的证据评估候选公理的可接受性的任务。对候选公理进行可靠评分的能力是自动模式或本体归纳所必需的,但它对本体和/或知识图验证也很有价值。准确的公理评分启发式算法通常在计算上很昂贵,如果您希望在迭代搜索技术中使用它们,例如逐级生成和测试或进化算法,这需要对大量候选公理进行评分,这是一个问题。我们解决了开发一个预测模型作为推理的替代品的问题,该模型预测候选类公理的可能性分数,并且足够快,可以在这种情况下使用。为此,我们使用从本体的包容结构中获取的语义相似性度量。我们表明,本工作中提供的方法可以准确地学习候选 OWL 类公理的可能性分数,并且可以对各种 OWL 类的公理进行学习。

#### 1142, 单调动态程序和应用程序的动态维护

Monika Henzinger, Stefan Neumann, Harald Räcke, Stefan Schmid, 2023.1.5

动态编程 (DP) 是算法设计的基本范式之一。然而,许多 DP 算法必须填写大型 DP 表,由二维数组表示,这至少会导致二次运行时间和空间的使用。这导致了对特殊情况下的改进算法的发展,当 DP 满足额外的属性,如 Monge 属性或完全单调性。

在本文中,我们考虑了一个新的条件,即假设 (除其他一些技术假设外) DP 表的

行是单调的。在这个假设下，我们引入了一种新的数据结构，用于计算  $(1+\epsilon)$  的近似 DP 解，在静态情况下，时间和空间接近线性，当 DP 条目动态变化时，更新时间为多项式。据我们所知，我们的新条件是以前的条件所无法比拟的，也是第一个允许在现有 DP 的基础上推导出动态算法的条件。我们没有使用二维数组来存储 DP 表，而是使用单调的分片常数函数来存储 DP 表的行。这使得我们只需使用  $\text{poly}(\log(n, W))$  位就可以存储长度为  $n$  的 DP 表行，并在  $\text{poly}(\log(n, W))$  时间内对这些函数进行操作，如  $(\min, +)$  卷积或舍入。

我们进一步介绍了我们数据结构的几个应用。对于双标准的  $k$  平衡图分区和同步源定位，我们获得了第一个具有次多项式更新时间的动态算法，以及第一个仅使用近线性时间和空间的静态算法。此外，我们获得了目前最快的全动态 knapsack 算法。

1143, 用于蛋白质复杂建模的 AlphaFold 基准测试揭示了准确性决定因素

Rui Yin、Brandon Y. Feng、Amitabh Varshney、Brian G. Pierce, 2022. 7. 13  
University of Maryland

蛋白质-蛋白质相互作用的高分辨率实验结构测定导致了有价值的机制见解，但由于大量相互作用和实验限制，需要能够准确模拟其结构的计算方法。在这里，我们探索使用最近开发的深度学习方法 AlphaFold 从序列中预测蛋白质复合物的结构。以 152 种不同的异二聚体蛋白质复合物为基准，测试了 AlphaFold 的多种实现和参数的准确性。值得注意的是，许多案例（43%）具有近乎原生的模型（对预测相互作用准确性的中等或高度关键评估）生成为 AlphaFold 的顶级预测，大大超过了未结合的蛋白质-蛋白质对接的性能（近乎 9% 的成功率） -

天然排名靠前的模型), 但是我们组中抗体-抗原复合物的 AlphaFold 建模是不成功的。我们确定了与缺乏 AlphaFold 成功相关的序列和结构特征, 我们还研究了多序列比对输入的影响。多聚体优化版本的 AlphaFold (AlphaFold-Multimer) 与一组最近发布的抗体-抗原结构的基准测试证实抗体-抗原复合物的成功率很低 (11% 成功), 我们发现 T 细胞受体-抗原该算法同样没有准确地模拟复合体, 表明适应性免疫识别对当前的 AlphaFold 算法和模型提出了挑战。总的来说, 我们的研究表明, 端到端深度学习可以准确地模拟许多瞬态蛋白质复合物, 并突出了未来发展的改进领域, 以可靠地模拟任何感兴趣的蛋白质-蛋白质相互作用。

#### 1144, 用于对话中情绪识别的上下文和情感感知网络

Geng Tu、Jintao Wen、Cheng Liu、Dazhi Jiang、Erik Cambria, 2022.2.7

Shantou University、Nanyang Technological University

对话中的情感识别 (ERC) 在推荐系统、人机交互和医疗保健等许多领域都具有广阔的应用前景。与其他情绪识别任务相比, 对话本质上是一个动态交互的过程, 在这个过程中, 人们经常依赖上下文和常识知识来传达情绪信息。本文提出了一个上下文和情感感知框架 Sentic GAT 以解决这一挑战。Sentic GAT 中的常识知识由基于情感一致性的上下文和情感感知图注意力机制动态表示, 上下文信息由具有层次多头注意力 (HMAT) 的对话转换器 (DT) 捕获, 其中 HMAT 是用于获取历史话语对自身和其他话语的依赖性, 以获得更好的上下文表示。本文探索了一种对比损失来区分情绪识别中的上下文无关和上下文敏感的话语, 以增强直接表达想法的直接对话中的上下文表示。实验结果表明, 上下文和情感信息可以促进常

识知识的表示，上下文话语的内部和相互依赖性有效地提高了 Sentic GAT 的性能，并且使用情绪强度的 Sentic GAT 在测试数据集上的表现优于最先进的模型。

#### 1145，人工智能的命运：走向算法包容性和可访问性

Isa Inuwa-Dutse, 2023. 1. 5

哈德斯菲尔德大学（英国）

这个时代的一个决定性现象是由人工智能（AI）技术驱动的系统广泛部署。随着人工智能占据中心舞台，社会的许多部门正受到算法决策的直接或间接影响。算法决策带来了经济和个人影响，这带来了公平、问责、透明度和道德（命运）的问题，在人工智能面向解决算法差异。伦理人工智能处理纳入道德行为，以避免在人工智能的决定编码偏见。然而，目前关于这些关键问题的讨论是由经济较发达国家主导的，这引起了人们对忽视当地知识、文化多元性和全球公平的关切。本研究建立在现有的关于负责任的人工智能的研究基础上，重点关注全球南方被认为相对于人工智能服务不足的地区。我们的目标有两个方面：（1）评估 FATE 相关问题和透明度方法的有效性；（2）提供有用的见解并激励采取行动，以弥合人工智能的可及性和包容性差距。使用来自在线社交网络的广告数据，我们设计了一个用户研究（n=43）来实现上述目标。研究的结果包括：对 AI 系统所做出的决定的解释往往是模糊的并且信息较少。为了弥合可及性和包容性之间的差距，有必要与社区合作，寻找将公平、问责、透明度和道德融入人工智能的最佳方式。这将有助于增强受影响社区或个人的能力，使其能够有效地调查和监管日益增长的人工智能驱动系统的应用。



敬请关注联盟微信公众号  
COPU开源联盟



扫描二维码  
获取往期资料

---

中国开源软件推进联盟秘书处

电话: +86 010-88558999

联盟公共邮箱: [office@copu.org.cn](mailto:office@copu.org.cn)

联盟官网: <http://www.copu.org.cn>

地址: 北京市海淀区紫竹院路66号赛迪大厦18层

---