



附件一：申报主题

1. 碳中和	2
1.1 面向数据中心算力需求和能源供给的智慧调度系统研究	2
2. 智慧医疗	2
2.1 医疗和生物信息中的先进人工智能技术研究	2
2.2 医疗 AI 领域中的知识驱动智能决策	3
2.3 医学多模态技术研究	3
3. 智慧交通	4
3.1 面向交通专网的高可用无状态核心网技术研究	4
3.2 基于视觉 SLAM 技术的建图和定位研究	4
3.3 基于人工智能的出行意图识别和推荐算法	5
3.4 面向智慧出行场景的视觉及多模态感知算法研究	5
3.5 基于大数据的交通分析、预测及优化	6
4. 机器学习与深度学习	6
4.1 不完备数据环境下的序列决策问题	6
4.2 GAME AI 之策略多样性研究	7
4.3 最优化出价策略在在线广告的应用	7
4.4 大数据分析 with 深度学习在攻击检测领域的应用研究	7
4.5 在线广告预估领域的多场景学习算法研究	8
4.6 面向复杂网络的可解释异常检测算法研究	8
4.7 AI 算法优化 EDA 及计算光刻学相关算法研究	9
5. 数字图像处理与计算机视觉	9
5.1 强鲁棒性的人体动作捕捉和重定向	9
5.2 视觉与多模态预训练模型研究	10
5.3 开放场景下跨 Domain 小样本物体检测在视觉场景的应用探索	10
5.4 全景图像及视频的理解与处理	11
5.5 实时人体表现捕捉与重建	11
5.6 工业视觉质检中的高效学习方法研究	12
6. 知识图谱与自然语言处理	12
6.1 自然语言生成	12
6.2 自监督的开放域对话持续学习	12
6.3 预训练模型微调方法与基于预训练模型的小样本学习	13
6.4 金融事件演变表示研究	13
7. 语音技术	14
7.1 多语言语音合成及跨语种音色迁移技术研究	14
8. 多模态融合	14
8.1 多媒体数字水印与视频内容篡改识别	14
8.2 基于深度学习的可调风格的短视频背景音乐合成	15
8.3 语音驱动的带表情 3D 面部动画合成	15
9. 软件工程	15
9.1 代码智能生成技术研究	15
9.2 面向二进制程序的软件安全关键技术研究	16
10. 密码学	16
10.1 不可信环境下的密钥保护研究与安全实现	16
11. 数据库	17
11.1 多模数据库的前沿研究	17
12. 网络技术	17
12.1 基于意图的下一代智能网络研究	17

申报主题介绍

(各主题均不限于给定的建议研究方向, 申请人可自行拓展决定。)

1. 碳中和

1.1 面向数据中心算力需求和能源供给的智慧调度系统研究

数据中心输入能源, 输出算力。传统数据中心输入能源(以市电为主)和设定输出算力都保持恒定, 数据中心运维重点是保持基础设施在额定容量和设计效率内可靠运行。碳中和战略将改变数据中心的能源利用范式, 对绿色能源和储能的使用比例逐步提升, 融合能源的供给不再恒定, 变为与时间相关的非线性参数, 特别是东数西算本质上要求根据数据中心输出算力需求动态匹配输入能源, 即数据中心运维目标转变为面向计算集群实现分布式源网荷储的动态调度。把数据中心想象为一台超级计算机, 智维平台就是其操作系统, 智维操作系统需要跨时空对数据中心集群进行统一接入、建模、监测、仿真和调度。本课题旨在建立起面向数据中心智能运维和能源调度的超域操作系统基础理论体系。

建议研究方向:

构建智维超域操作系统基础架构理论体系, 包括硬件抽象层、数据中台层、应用界面层、算法平台层和应用商店层。重点面向碳中和及东数西算场景, 在智维操作系统基础上研究数据中心微电网和储能综合调度策略; 研究元宇宙发展对数据中心运营维护和数据架构的影响; 推论智维操作系统在泛工业行业的应用前景, 对数字新基建架构和数据中心行业标准的影响力, 构建开放和共赢的生态环境。

[返回目录](#)

2. 智慧医疗

2.1 医疗和生物信息中的先进人工智能技术研究

基于机器学习的人工智能技术已经在诸如蛋白质结构预测、新药发现、疾病诊断、个性化治疗规划和预后预测等富有挑战性的前沿生物医学领域取得了显著成果。鉴于生物数据和医疗数据的特殊性, 需要设计能够有效融合领域知识的人工智能模型算法。此外, 医学和生物学数据正变得越来越多模态, 这些多种不同模态的信息综合反映了生物过程之间复杂但丰富的潜在关系。本课题旨在研究和开发用于医疗和生物信息数据的分析、挖掘和理解的先进人工智能技术。

建议研究方向:

- 1) 免疫组库分析中的前沿人工智能技术, 包括但不限于抗体-抗原亲和力预测, T/B 细胞受体抗原特异性预测, 疫苗有效性分析等;
- 2) 用于空间转录组学、单细胞基因组学和转录组学的前沿人工智能和多模态学习技术;
- 3) 用于精准医疗的多模态/多组学融合技术和数据, 包括但不限于基于多模态融合的疾病诊断、治疗干预和病人预后预测。

[返回目录](#)

2.2 医疗 AI 领域中的知识驱动智能决策

尽管人工智能已经在医疗领域证明了其有效性，但临床医生依然难以在实际工作中使用这些 AI 工具。其中很重要的一个原因是这些 AI 工具所做出的决策缺乏足够的可解释性。本课题旨在研究医疗 AI 领域中的知识驱动智能决策，帮助有效解决以上问题并同时提高 AI 算法的透明度。

建议研究方向：

- 1) 医疗知识图谱构建，该方向旨在构建一个有效并符合医学专业认识的知识图谱，该知识图谱需要对复杂的医学知识进行合理的刻画。无论是否基于三元组格式构建，该知识图谱需要兼容适配现有的 AI 算法，如 KBQA 知识问答算法；
- 2) 医疗信息抽取，该方向旨在研究医疗实体与医疗事件识别算法，该算法需要支持不同医学术语的识别（例如疾病或药物）；支持不同场景领域的识别（例如电子病历或医疗对话场景）；支持不同任务设定下的识别（例如低资源设定）；
- 3) 基于知识的疾病预测，该方向旨在利用多模态及时序患者数据，给出具有可解释性的诊断或治疗预测结果。同时该方向也强调对于可解释性的探索，包括全局/局部可解释性、后置/前置可解释性、可解释性的量化。

[返回目录](#)

2.3 医学多模态技术研究

医学服务的数字化产生了大量的多模态数据，包括文本、图片、影像、时序数据等。这些数据蕴含着大量的知识，而目前没有被很好的挖掘利用。本课题旨在探索如何深入理解多模态数据，挖掘出可用知识，更好的服务于患者和医生的相关研究。

建议研究方向：

- 1) 无结构文本中的医学知识抽取、表示和推理；
- 2) 结合医学知识图谱，构建医学文本预训练模型，对于特定场景，也可以尝试融合文本和医学图片的多模态预训练模型；
- 3) 医学深度学习算法可解释性、公平性及如何保护隐私方面的研究；
- 4) 针对数据稀疏和分散的医疗场景下的无监督、半监督、迁移学习和元学习等算法研究，包括但不限于：通用模型在罕见病上的适用，数据孤岛条件下的模型训练，少样本上的疾病模型训练等；
- 5) 集合多模态数据患者诊疗周期中的预测服务研究（疾病阶段预测、恶化预测、复发预测等）。

[返回目录](#)

3. 智慧交通

3.1 面向交通专网的高可用无状态核心网技术研究

智慧交通中的交通专网将承载海量的车辆感知、控制和数据分发等业务，需要保障网络的高可用性。由于车联网所承载的数据具有时空不均衡性，交通专网资源需要针对业务波动性进行灵活的水平伸缩。但当前 5G 核心网的部署通常是状态与协议处理紧耦合，核心网各网元难以进行灵活伸缩，网络扩缩容或部署网元出现故障时将会导致用户终端业务受影响。无状态技术能够在专网核心网充分利用云环境下分布式资源特性，各网元服务以多实例服务网格形式提供，从而保障云环境下的业务高可用行、灵活伸缩能力。然而目前 5G 核心网的无状态化技术仍处于起步阶段，尚没有关键性能分析及原型系统验证。

本课题旨在研究核心网关键网元的状态与协议的解耦机制，实现核心网网元状态的统一存储；研究各实例间信令级负载均衡算法，实现针对业务及资源分布的信令调度，保障核心网的高可用性，形成无状态化 5G 核心网原型系统。课题组提供 5G 云化核心网软件环境。

建议研究方向：

- 1) 核心网关键网元状态及协议及其负载均衡实现机制。针对核心网控制面和用户面网元所使用的 SCTP、HTTP2.0、GTP 等协议，结合 UE 状态控制、PDU 会话状态和用户面业务数据会话状态，定向研究其负载均衡的实现机制；
- 2) 网元实例化后的负载均衡算法研究。针对核心网控制面和用户面网元所使用的 SCTP、HTTP2.0、GTP 等协议，结合协议 IP 包体关键业务数据字段，使得负载均衡设备可以根据业务场景对 RAN 与 5G 核心网控制面、RAN 与 5G 核心网用户面、5G 核心网网元之间的请求进行灵活高效的负载均衡处理；
- 3) 结合网元状态和协议分析的业务处理和业务状态数据解耦机制。针对 UE 移动性管理 MM 状态和会话管理状态 SM 状态，通过改造业务处理机，实现业务处理和业务会话数据的解耦，进一步实现无状态的业务处理机集群，在单个业务处理机故障情况下，可无缝切换到其它的业务处理机，保证业务会话的连续性；
- 4) 网元状态统一存储机制及实现。基于内存和数据库同步技术的业务会话状态数据存储和获取，实现业务会话状态数据的持久化存储。当业务处理机集群中的某一节点宕机，业务会话被集群内其它节点接管情况下，业务会话状态数据可实时获取，从而保证业务会话处理的连续性。

[返回目录](#)

3.2 基于视觉 SLAM 技术的建图和定位研究

视觉 SLAM 作为机器人自主导航的经典方案，获得了越来越多的关注，其要求简单、成本低廉、易于实现，在研究深度和创新性方面也不断提升。在封闭场景内，该技术可以根据视觉采集数据完成基于点云的三维地图重建，进而可基于点云地图进行精准定位，解决封闭场景内卫星信号无法定位的问题，补齐在封闭空间高精定位和持续导航的能力短板。在室外大场景，视觉 SLAM 技术同样可以提升定位精度，并在道路语义层面进行建模，构建道路语义地图。但目前视觉 SLAM 仍旧面临诸多挑战，如场景变化、纹理缺失、运行效率、尺度恢

复等问题。本课题旨在结合腾讯地图在各种场景高精度定位和建图的需求，探索和实现高精度定位和建图应用。

建议研究方向：

- 1) 研究和探索面向封闭场景和室外开阔场景的视觉 SLAM 新技术和新应用；
- 2) 开展对已有 SLAM 框架、语义 VIO 等技术方案的效果对比，研究性能提升方法，探索视觉 SLAM 在三维点云重建和精准定位的应用潜力。

[返回目录](#)

3.3 基于人工智能的出行意图识别和推荐算法

越来越多的用户在地图产品中进行驾车导航、POI 检索、路况查询、路线规划，产生了大量的出行数据。基于这些数据有效识别用户出行意图、预测出行需求，利用推荐算法为用户提供高效的出行助手服务，成为一个非常有商业价值和研究价值的课题。本课题旨在研究基于人工智能的出行意图识别和出行类服务推荐技术，包括用户画像建模、意图识别、推荐算法等。

建议研究方向：

- 1) 用户画像建模技术，包括 user embedding、图算法等；
- 2) 推荐技术，包括 CTR 预估、多目标建模等。

[返回目录](#)

3.4 面向智慧出行场景的视觉及多模态感知算法研究

智慧出行涵盖智能座舱、自动驾驶等多个场景，也面向高、中、低端多层次硬件平台。感知是出行领域不可或缺的部分，而针对上述多场景、多硬件平台提供适配性、一体化的感知方案具有不小的挑战。面向中低端座舱平台，多任务感知架构能够极致化地利用有限的算力，为座舱下游服务的智能化提供有力的支撑，如何在此基础上进一步进行多任务协同、单任务定点优化是我们不断探索的方向。与此同时，视觉单模态已无法承载快速演变的座舱需求，视觉、语音、NLP 跨模态的融合是当下比较明晰的发展路径。而面向高算力云端平台，高度依赖感知的数字化场景提取和重建也是自动驾驶工具链中重要的一环。本课题旨在探索基于海量的计算资源和标注数据，打造更高性能的感知引擎来提升出行体验的方法。

建议研究方向：

- 1) 多任务学习协同优化：多任务学习能实现有限算力下更高的检测精度和更广的检测要素类别。多任务协同优化包括但不限于：多任务架构的设计和组合、异源数据之间的调度机制以及任务间损失自适应调节等；
- 2) 多模态环境感知与场景理解：通过模型前端融合视觉环境感知、语音识别及语义理解等跨模态任务，实现端到端的场景理解；
- 3) 高保真数字化场景提取及重建：结合 Camera、Lidar、IMU 等传感器数据进行多源环境感知及三维重建。

[返回目录](#)

3.5 基于大数据的交通分析、预测及优化

高效绿色的城市交通需要精准感知预测交通动态、深度挖掘出行规律和问题、优化规划和管理、引导需求和服务等。本课题旨在通过对海量交通数据和城市动态数据的深度分析，从多尺度、多角度建立科学系统的交通评估、诊断、预测及优化方法体系，融合网络动态、宏观规律、微观行为、出行需求、绿色低碳、便利可达等多源信息和维度，实现交通精准感知、预测、服务和管理，支持城市深度理论分析和决策优化。课题组提供脱敏数据和落地场景。

建议研究方向：

- 1) 科学系统的城市体检和交通评估诊断指标体系和计算方法；
- 2) 实时交通感知、分析、预测和仿真：比如动态出行需求感知、交通规律挖掘和可视化、车流和交通状态预测、微观和宏观交通仿真；
- 3) 交通优化和服务：比如公交线路优化、混合交通模式的路线规划、大规模个性化导航路径规划技术和服务模式等；
- 4) 绿色出行：通过对城市交通设施、路网结构、出行需求、交通方式选择、可达性等综合分析，引导绿色出行，推动城市交通规划、管理和生活方式的绿色转变、高质量发展和公平服务。

[返回目录](#)

4. 机器学习与深度学习

4.1 不完备数据环境下的序列决策问题

从仿真器或者高质量离线数据中，强化学习可以学到优秀的策略以实现自动化控制。然而，在很多真实应用中（如温室控制），我们难以获得仿真器或者高质量离线数据，而仅仅可以获得不完备的数据环境。在不完备数据环境下，如何提升强化学习的各种性能，例如稳健性、有效性等，是本课题的主要研究内容。

建议研究方向：

- 1) 基于低质量的离线数据，研究稳健的离线强化学习算法；
- 2) 基于多个不完备数据环境，研究迁移强化学习、以及元强化学习；
- 3) 针对马尔可夫假设不成立的情况，研究相应的强化学习算法。

[返回目录](#)

4.2 GAME AI 之策略多样性研究

“多样性”本身作为一个研究方向，广泛存在于诸如计算机科学、生物学、经济学等各个学科中。近年来，策略多样性作为 GAME AI 里的一个研究方向越来越受到学界重视。一方面，出于对“多样性”本身的需求和好奇，策略多样性提供了对同一问题的不同解决方案。另一方面，策略多样性作为一个模块，对其他问题的解决也发挥着至关重要的作用。同时，在工业界，随着 AI 在游戏里的广泛应用，用户对 AI 的智能化和多样性都提出了更高的要求。本项研究旨在研究 GAME AI 领域策略多样性生成的相关前沿技术，同时探索、落地策略多样性在 GAME AI 领域的应用价值。

建议研究方向：

- 1) 策略多样性的自动生成：在相应游戏场景中，根据用户定义的多样化指标，生成多样化的策略池，其中每个策略本身满足一定的质量约束（质量约束是指对 AI 强度的要求）；
- 2) 策略多样性应用：探索多样化策略池对相关问题的应用价值，相关问题包括但不限于：GAME AI 场景中的人机协作、对手/队友建模；游戏本身的多样性和可玩性；游戏 NPC 的智能性和多样性等。

[返回目录](#)

4.3 最优化出价策略在在线广告的应用

计算广告是当前工业界算法应用最为成熟的领域之一，游戏展示广告是各大广告平台重要的业务组成部分。虽然实时竞价策略可以根据模型对每个流量的预测指标（比如预测的点击率、转化率、用户价值等）评估最终出价，但如何考虑市场因素、整合不同的评估指标以更高效地给出最优出价，仍然存在诸多技术挑战，如市场价格预估、冷启动出价等。本课题旨在基于腾讯游戏海量脱敏数据，利用强化学习等前沿算法框架，最优化游戏展示广告场景下实时竞价广告的出价策略，提升出价效率。

建议研究方向：

- 1) 设计高效且实用的多场景、动态化的强化学习算法与框架，支持序列决策，优化长期收益；
- 2) 利用强化学习算法结合实际业务问题，改进、优化当前出价方式，提升整体消耗；
- 3) 基于强化学习的可拓展性，探索优化冷启动广告出价问题的方法。

[返回目录](#)

4.4 大数据分析 with 深度学习在攻击检测领域的应用研究

随着信息系统复杂度的增大以及攻击手段的不断演进，传统静态规则检测方案过于依赖人工经验，且很容易被攻击者绕过。本课题旨在把大数据分析和深度学习相关技术（如数据挖掘、自然语言处理、图神经网络等）应用于网络攻击检测研究中，其成果可以弥补静态检测规则过于依赖人工经验的缺点，从数据中挖掘更多潜在的攻击。

建议研究方向：

- 1) 基于进程树语义表征的恶意属性分析研究，例如评估进程的威胁等级、攻击分类；
- 2) 系统命令的表征和分类研究，例如识别同样功能的不同表示、命令主成分信息；
- 3) 进程树结构匹配问题研究，例如评估完整进程树的相似度、评估局部进程树的相似度、对相似程度打分；
- 4) 基于攻击过程的特征收集攻击线索，结合 ATT&CK 攻击阶段绘制新攻击的攻击路径；
- 5) 良性主机的进程白名单方案研究，例如进程白名单防误报、过滤进程白名单后告警。

[返回目录](#)

4.5 在线广告预估领域的多场景学习算法研究

工业界的在线广告系统中，点击率、转化率等精排模型的预估是核心的技术。在联盟广告的场景下，由于不同的接入方式、诸多的广告位样式、众多的媒体，精排模型的预估效果存在诸多技术挑战，集中表现在不同场景下的偏差。而经典的多场景学习算法，在面对联盟广告复杂的场景问题时，参数量、分场景样本量、梯度更新等问题尤为突出，影响模型效果。本课题旨在基于腾讯联盟广告相关脱敏数据，提升不同场景下的精排模型预估效果，尤其是研究多场景学习算法，并迁移应用到联盟广告的预估场景中。

建议研究方向：

- 1) 调研包括但不限于 MMOE、ESMM、SNR、CGC/PLE、STAR 等广告推荐场景下常用的多任务学习算法，结合业务做算法适配及优化创新，应用到广告精排点击率、转化率预估的场景；
- 2) 以算法效果和落地效率为出发点，思考从模型结构、损失函数、优化器、样本构造等角度，设计工业界普遍适用的多场景学习算法，并通过 AB 实验将算法在多个业务场景落地；
- 3) 建立合理的多场景下的精排预估模型离线效果评估指标；建立标准的多场景下的精排预估模型的训练、测试样本集。

[返回目录](#)

4.6 面向复杂网络的可解释异常检测算法研究

近年来深度学习逐渐在异常检测领域得到重视与应用，尤其在时序数据、图数据上的异常检测研究取得了很大的进展。然而随着模型复杂度的提升，可解释性与稳定性逐渐成为这类算法面临的新挑战，同时可解释性不足也为这些模型的大规模业务应用增加了难度。当前对于深度学习可解释性的研究主要集中在以下方向：1) 对模型的参数或者输入施加调整或者扰动，对模型最终结果的变化进行分析，以此推断模型作出决策的依据；2) 直接构建本身就具有可解释性的模型或者模型子结构，或是对已有的复杂模型构建具有可解释性的代理模型。这些研究方向已经取得了一定进展，然而在许多具体的领域应用中，仍然需要比较多的人工调整和针对性优化。本课题旨在能够结合产业真实场景特性，构建出兼顾可解释性与效果的异常检测模型。

建议研究方向：

- 1) 构建面向大规模支付网络的动态图异常检测模型；
- 2) 使用隐层分析、因果推理或引入可解释性子结构对上述模型的决策输出可解释性的证据。

[返回目录](#)

4.7 AI 算法优化 EDA 及计算光刻学相关算法研究

近年来，多种机器学习算法被广泛应用于 EDA 相关的算法场景中，从版图布局布线，到计算光刻学的相关算法，都受益于通过引入机器学习算法而获得的运行效率提升。对于这个领域的研究大致可以分为两个方面：一个是大数据及机器学习的应用对于算法本身的加速，另一方面是用算法（如深度学习）来替代传统的物理场仿真，而达到提升运行效率的目的。本课题旨在通过以上两个方向的尝试，进一步挖掘和探索提升 EDA 底层算法效率的空间。

建议研究方向：

- 1) 研究用深度学习等 AI 算法代替光学等物理场仿真，优化计算光刻学中的仿真算法；
- 2) 用机器学习方法加速 EDA 算法，包括但不限于布局和布线以及物理验证等相关问题。

[返回目录](#)

5. 数字图像处理与计算机视觉

5.1 强鲁棒性的人体动作捕捉和重定向

基于视觉的动作捕捉是一个热门的研究领域，在增强/虚拟现实等应用中有突出的应用前景。然而，这项技术面临两个挑战：一是现有的动作捕捉算法在鲁棒性和实时性方面都存在不足；二是缺乏成熟的动作重定向算法，使得捕捉到的动作无法驱动任何虚拟角色。本课题旨在研究准确、稳健、高效的人体姿态估计和运动重定向算法，为数字虚拟人等应用提供关键技术支撑。

建议研究方向：

- 1) 3D 人体姿态估计：如何利用时序信息、运动先验等信息提升模型的鲁棒性以及如何利用知识蒸馏等方法对模型进行轻量化；
- 2) 3D 全身姿态估计：如何利用面部、手部、肢体的特性及其空间关系等信息，提升全身姿态估计算法的精确度；
- 3) 动作重定向中的差异结构、自接触与自穿模：利用深度学习解决人体结构存在差异的两个主体间的动作重定向问题，同时保留源人物的自接触并避免目标人物的自穿模。

[返回目录](#)

5.2 视觉与多模态预训练模型研究

随着模型规模的不断增大，对训练数据的要求也成倍增加，自监督预训练模型成为当前学术界与工业界的研究热点，并且有了越来越多的实际落地应用。这其中即包括视觉、文本等单模态预训练，也包括混合多模态预训练。本课题旨在对视觉预训练、多模态预训练进行研究，探索模型结构、自监督任务的设计等，以提升预训练模型效果。研究成果需要在业界公开的测评任务（如跨模态检索任务、分类任务、视觉常识推理任务等）上达到 SOTA 水平，并在该领域的国际顶会或者期刊上发表论文。

建议研究方向：

针对给定的上亿级别的多模态脱敏数据，包括但不限于从以下几个方面展开研究：

- 1) 当前主流的多模态预训练中，视觉特征的表达通常采用 region feature、grid feature、patch feature 中的一种，如何设计一个更好的视觉预训练模型，能够提取出一个兼具上述 3 种 feature 特性的表达，或者如何在多模态预训练中，将上述 3 种 feature 进行更好的模型结构设计融合；
- 2) 当前的自监督预训练方法往往针对通用任务设计，如常规的 MIM/MLM/ITM 等；如何针对不同的下游目标任务，如分类/检测/检索/VQA/生成等，增加设计一些有针对性的模型结构和自监督任务，从而进一步提升各目标任务效果；
- 3) 针对业务场景中数据模态多变（文本、图片、视频）的情况，设计一种兼容单模态、部分多模态、全模态的统一混合模型结构和训练方法。

[返回目录](#)

5.3 开放场景下跨 Domain 小样本物体检测在视觉场景的应用探索

物体检测技术已经在工业界有广泛应用，但实际场景中需要的检测目标类别随着应用需求不断更新变化，在实时性高的场景需求中迭代尤为迅速。针对类别更新而进行的常规模型迭代流程包括收集大量数据、数据标注、模型迭代训练直至满足上线的性能要求。整个迭代过程冗繁漫长，无法应对快速的类别更新需求。针对该问题，小样本检测提供了一种快速且有效的解决思路。本课题旨在依托海量无标签数据预训练底层模型，并利用少量的几个或者十几个标注样本，在不需要训练或者几分钟训练的情况下快速更新模型以适应新增的物体类别，满足信息产业中无感知更新的快速迭代要求。现有小样本检测方法试验的数据集多是限定场景，没有跨 Domain 的开放场景的应用能力。此外，视频数据上的连续信息在小样本检测上也未被充分挖掘。课题组提供相应的实验环境、训练资源和脱敏 Wild dataset 数据。

建议研究方向：

- 1) 信息产业跨 Domain 的开放场景下小样本目标检测与自监督方法的结合；
- 2) 信息产业跨 Domain 的开放场景下元学习和传统检测训练在小样本新类别增量上的结合探索；
- 3) 信息产业跨 Domain 的开放场景下小样本目标检测的模型轻量化和快速化，支持亿级数据快速检测；
- 4) 视频数据上小样本物体检测算法研究。

[返回目录](#)

5.4 全景图像及视频的理解与处理

近年来，沉浸式、全景媒体方面的研究及行业应用已引起广泛关注。与传统 RGB 图像、视频相比，360 全景媒体具有 $360^\circ \times 180^\circ$ 的视场观测角度；同时，对于 360 全景媒体，观测者可以选取任意角度进行自由观看，因此 360 全景媒体具有更强的交互性，并且可以给予观测者更真实的观感与沉浸式体验。相关技术有望被广泛应用于 VR/AR、机器人、自动驾驶，以及元宇宙构建等领域。

目前对于 360 全景媒体的研究基本上都是基于其在普通视场角二维平面上的投影来进行，而基于 360 全景媒体球面成像的原理，图像信息会在平面投影上产生失真，特别是在两极位置。在将现有基于一般平面媒体的技术与算法应用于全景媒体时，往往会产生明显的效果折损。因此，针对全景媒体的处理、理解任务（如语义分割、超分辨率、深度估计等），需要结合图形学、计算几何学及深度学习等方法，从整体上进行系统性设计。

本课题旨在探索全景媒体处理、理解相关的研究技术及方法。

建议研究方向：

- 1) 针对全景图像或视频的半监督/弱监督超分辨率算法；
- 2) 全景图像或视频的深度估计与深度超分辨率算法；
- 3) 基于全景图像或视频的三维场景结构理解及生成。

[返回目录](#)

5.5 实时人体表现捕捉与重建

随着元宇宙概念的推广，如何对现实中人物的形态及动作进行数字化，并在虚拟世界中重现、进行社交互动是当前学术界与业界热门及前沿的研究领域，该项技术在广电、娱乐、体育、健康等领域都着巨大应用前景。本课题旨在通过计算机视觉或图形学的算法研发，实现高品质的实时人体表现的捕捉和重现，这里的“表现”并非特指姿态，而是体型、外貌、表情、手势中的一样或多样。对于具体的技术路线则并没有特定限制，无论是单目或多目方案均可；无论是偏传统的 MVS，人体参数化模型或近年流行的深度隐式表达、神经辐射场（Nerf）均可。课题组可以基于云游戏的基础技术架构，提供多视角低延迟视频的拍摄、同步、上下行传输、云端渲染、云端计算方案。

建议研究方向：

- 1) 基于单目/多目 RGB 信息的高品质人体 3D 建模；
- 2) 基于姿态的人体重建与生成；
- 3) 虚拟视点合成、自由视角渲染技术（Virtual view synthesis, free-viewpoint rendering）。

[返回目录](#)

5.6 工业视觉质检中的高效学习方法研究

智能制造国家战略下，制造业正在加速向数字化、智能化转型。精密制造过程对高效高质的缺陷质检需求强烈，使用前沿的 AI 视觉质检技术取代传统的人工质检已势在必行。另外，与此相关的工业异常缺陷检测技术，以及对标注数据依赖性较低的高效学习方法也逐渐成为研究热点。为此，本课题旨在针对工业视觉质检场景，研究异常检测技术及低标注需求下的半监督、弱监督、缺陷挖掘与生成、域适应等技术。

建议研究方向：

- 1) 研究工业质检中的异常检测；
- 2) 研究工业质检中的半监督和弱监督学习；
- 3) 研究工业罕见形态的缺陷挖掘与生成技术；
- 4) 研究工业质检中的增量学习、主动学习、在线学习等模型迭代技术；
- 5) 研究工业质检中的迁移学习与领域自适应。

[返回目录](#)

6. 知识图谱与自然语言处理

6.1 自然语言生成

自然语言生成是指在特定的交互目标下，根据给定的输入信息生成人类可读的自然语言文本的自动化过程。近几年来，本领域的技术取得了重大的突破和进步，在媒体、娱乐、教育、电商等多个行业中均展现出巨大的应用潜力。目前，自然语言生成已经成为自然语言处理乃至人工智能领域的热门前沿方向，受到学术界与工业界的广泛关注。本课题旨在探索自然语言生成相关的技术和方法。

建议研究方向：

- 1) 大型语言模型和其他生成模型，包括新颖的模型架构和高效的训练方法；
- 2) 长文本生成，例如故事生成和文档生成；
- 3) 可控的文本生成，包括以属性、提示、表格、检索语句、图像或视频为条件的文本生成；
- 4) 知识、语义和其他推理信息在自然语言生成中的应用；
- 5) 生成模型的模型分析，包括可解释性分析、鲁棒性分析、攻防分析；
- 6) 开放式生成任务的自动评估方法。

[返回目录](#)

6.2 自监督的开放域对话持续学习

构建可持续学习的开放域对话模型是极具挑战性的任务。首先，开放域对话数据涉及多个领域，现有模型很难同时从多领域数据学到有效知识；其次，现有模型往往依赖大量标注数据，存在标注数据花销大、工作模式不可持续等问题；最后，开放域对话系统如何利用外

部知识进行自监督学习框架探索仍然缺乏深度研究，包括通用对话策略、对话状态、知识体系的定义、设计与更新等。

本课题旨在研究如何构建可持续学习的开放域对话模型，在充分利用已有多领域数据基础上，实现自监督地学习新领域知识，完成模型自我知识的持续更新。

建议研究方向：

- 1) 如何从多领域对话数据上做到有效知识迁移，从而学习到统一有效知识；
- 2) 在保证不忘记已有能力的前提下，如何学习新任务；
- 3) 如何通过自监督和持续学习完成对知识体系与用户特性的构建和更新；
- 4) 如何定义通用的基于符号的开放域对话状态，以及对话过程中如何动态更新对话状态；
- 5) 如何定义通用的对话 Action，以及如何自动学习对话策略。

[返回目录](#)

6.3 预训练模型微调方法与基于预训练模型的小样本学习

最近，大规模预训练语言模型通过大量无监督语料以及合理的自监督任务提升了文本表征能力。在小样本学习的场景中，由于训练数据稀少，提升模型的泛化能力以及迁移能力成为关键。一系列研究表明，预训练语言模型积累了语言知识，并且具备一定的知识推理能力，从而有效提升了原有深度模型的泛化能力。借助预训练模型，通过预训练与微调的范式，在多个场景的小样本学习中均取得了不错的效果。本课题旨在探索预训练模型微调范式，进一步消除预训练和微调阶段的差异，更好的利用预训练模型中丰富的语义知识，从而提升小样本学习效果。

建议研究方向：

- 1) 改进微调范式，分析主流微调范式 (Prompt-tuning, adapter) 和不同下游任务之间的差异，提出更贴合下游任务的微调范式；
- 2) 数据增强，利用下游任务中大量的无监督数据，提升微调阶段的效果。

[返回目录](#)

6.4 金融事件演变表示研究

金融资讯包括财经新闻、财报、研报、社区评论等，对股价预测、投资风险控制有着重要影响。资讯中的金融事件通过相同实体、相关行业概念、同类型事件、不同事件的时序或因果关系等构成了一个复杂且不断演变的网络。新的金融事件对不同股票的价格影响不一，相同类型的事件在不同时间产生的影响也不尽相同。本课题旨在研究如何表征金融事件的动态演变，进而提取有效的股价预测因子或投资风险特征，是人工智能赋能量化投资、金融风险控制的重要课题。

建议研究方向：

- 1) 金融舆情事件的网络构建与更新技术；
- 2) 个股、行业板块的金融事件演变特征提取；

- 3) 基于不同时间粒度的金融事件股价预测因子构建;
- 4) 基于不同时间粒度的金融事件投资风险识别特征构建;
- 5) 效果验证方案设计与评估。

[返回目录](#)

7. 语音技术

7.1 多语言语音合成及跨语种音色迁移技术研究

近些年来端到端语音合成技术已大大提升合成语音的表现力及自然度,然而如何构建一个通用的高自然度的多语言语音合成系统仍具挑战性。通常在多语言语音合成模型训练中,仅含有来自不同语言的单语说话人数据可用(例如数据中含有说话人 A 的中文语音数据,说话人 B 的英文语音数据,说话人 C 的日文语音数据,但来自同一个说话人的中/英/日跨语言数据非常稀缺)。在不引入额外训练策略和方法的情况下,仅基于来自单语说话人的多语言数据训练的合成模型在进行跨语言合成时往往具有目标说话人音色不一致或自然度较低的问题。此课题旨在提出更好的模型结构以及引入更好的训练策略允许在多语言语种间进行说话人音色、风格等迁移实现合成语音在语种、音色、风格的自由组合。

建议研究方向:

- 1) 利用 domain adversarial 学习对声学模型编码后表征进行语言,说话人,内容等因子进行解耦;
- 2) 模型中引入合适结构的说话人和语言编码器,利用多任务学习在两个编码器分别施加说话人分类和语言分类损失函数获取说话人和语言的信息同时引入正交损失鼓励两个编码器的解耦等。

[返回目录](#)

8. 多模态融合

8.1 多媒体数字水印与视频内容篡改识别

随着多媒体技术和网络通信的发展,数字媒体的安全隐患日益严重,一方面平台需要有能保障创作者权益的能力,另一方面平台需要有能识别数字媒体是否被篡改的能力。本课题旨在探索保障多媒体数据安全的技术和方法,提升平台的安全等级,保障互联网生态安全。

建议研究方向:

- 1) 多媒体数字水印技术,研究方向不限于图像/音频/文本的数字水印算法建模和实现;
- 2) 视频内容取证/篡改识别的算法研究工作,包括基于元数据分析、深度学习方法建模,用于拼接检测以及篡改定位。

[返回目录](#)

8.2 基于深度学习的可调风格的短视频背景音乐合成

随着短视频的风靡，我们的娱乐方式和信息获取的方式有了很大的改变。合适风格的背景音乐可以增强短视频的情感氛围以及表现力，并显著提升用户的体验和情感反应。然而，背景音乐高昂的版权费和有限的版权使用时间使得短视频制作成本变得很高，并且选取符合短视频风格的音乐也需要一定的专业创作门槛。现有的音乐智能生成技术生成的音乐结构比较简单，其情感、风格难以符合短视频的整体需求。本课题旨在研究如何使用深度学习技术生成结构、情感饱满，并与短视频风格相匹配的背景音乐，从而降低短视频的制作成本和创作门槛。课题组可以提供附有背景音乐的短视频训练样本，具有版权的音乐数据集和所需的计算资源。

建议研究方向：

- 1) 音乐结构的生成模型：探究利用生成对抗网络、变分自动编码器等深度生成模型，在给定的视频风格的嵌入下，生成符合特定风格的音乐；
- 2) 多乐器的复杂音乐的合成：探究使用生成模型在多乐器条件下合成出使用这些乐器创作出的高质量旋律或伴奏；
- 3) 视频风格提取：探究使用深度网络，提取视频的风格信息，用以作为风格条件合成音乐。

[返回目录](#)

8.3 语音驱动的带表情 3D 面部动画合成

语音驱动的 3D 口型动画近些年在学术界得到了广泛的研究，并取得了不错的成果，同时在游戏动画制作中也获得了广泛的应用。但是，通过用户指定情绪或由语音直接合成带有情绪的、逼真的面部动画，仍是未解决的难题。本课题旨在实现一种由语音驱动的、带表情的面部动画合成方案，目标为：1. 能够根据语音生成自然且精确的唇形动画；2. 能够根据用户指定的情绪类型及其强度，自动合成出逼真的面部动画(包括面部、嘴部、舌头、及头部的运动)；3. 能够从语音直接提取出用户情绪，与口型动画相结合，合成自然的、带表情的面部动画。

建议研究方向：

采用监督学习方法，从多模态数据库中（包含同步的语音、文本、面部动画等数据）学习，来合成自然的、带表情的 3D 面部动画，情绪不仅可以由语音或文本智能提取，也可以通过用户直接指定。

[返回目录](#)

9. 软件工程

9.1 代码智能生成技术研究

由于产品和业务逻辑的复杂性增加，当前软件开发的规模和复杂度也在不断的增加。如何高效的研发大规模、复杂特性的软件，提高软件开发的效率是当前亟待解决的一个问题。

题。因此，代码生成技术得到重视。代码生成是机器根据结构化或非结构化的功能描述自动生成源代码的过程。传统上已有基于经验、模板、规则的方式生成的代码，其一般受限于领域特性，生成的代码较为局部，难以获得通用性。随着深度学习、Transformer 模型等的飞速发展，如何采用机器学习的方式从代码大数据中学习从而指导机器自动生成代码是当前一个热门研究方向。本课题旨在探索基于代码描述或注释等信息自动生成代码的技术，或者基于代码上下文的自动补全技术，以提高软件开发的效率。

建议研究方向：

- 1) 基于函数/类/文件描述信息或注释的代码生成技术；
- 2) 基于程序上下文的代码行/代码块补全技术；
- 3) 特定场景的代码生成技术。

[返回目录](#)

9.2 面向二进制程序的软件安全关键技术研究

随着软件复杂度的不断提升，面向二进制程序的分析与安全保障工作面临新的机遇和挑战。本课题旨在结合学术界研究热点，选取真实业务场景密切相关的研究课题，研究成果可以对二进制程序分析中的逆向工程、漏洞挖掘等关键任务产生积极影响。

建议研究方向：

- 1) 机器学习技术在二进制程序分析场景下的应用研究，包括但不限于二进制代码相似度分析、无符号二进制文件的符号补全、二进制代码与源代码的匹配与翻译；
- 2) 供应链安全场景下面向二进制程序的分析技术研究，包括但不限于二进制可执行文件的软件成分分析，如第三方库及其版本号等关键信息识别、已知漏洞的识别检索；
- 3) 基于抽象解释的二进制程序静态分析技术研究，包括但不限于值集分析、抽象解释下的安全属性研究、降低误报方法研究；
- 4) 编译器（解释器）测试代码的自动生成和优化技术研究，包括但不限于自动生成用于编译器（解释器）模糊测试符合特定程序设计语法结构的程序代码及其预期结果、面向上述测试的过程优化技术。

[返回目录](#)

10. 密码学

10.1 不可信环境下的密钥保护研究与安全实现

不可信环境下的密钥保护一直是工业界的一大课题，密码安全控件实现的安全高效应用成为密码合规的重要一环，具有重要的研究意义。然而，密钥保护模块也面临很多业界难点需要攻关。首先，在安全性方面面临很大的挑战，如：不可信环境下的内存安全风险、终端设备的侧信道攻击风险、随机数安全等；其次，硬件设备、操作系统种类繁多，存在着各种各样设备兼容、系统兼容问题，如何设计出一种通用、安全、可信的密钥保护方案还需

学术界与工业界共同努力。本课题旨在结合现有密码算法与技术方案，在密钥保护模块的安全性和工程实现上进行探索，攻克业界难题，联合推动密钥保护相关技术领域标准的制定。

建议研究方向：

- 1) 研究可增强密钥安全性的方法，如 TEE 可信执行环境、安全虚拟机技术、内存安全域、白盒加密等；
- 2) 实现不可信环境下的密钥保护解决方案，包括但不限于工程实现、技术文档、安全论证等；
- 3) 结合密码安全控件应用场景，设计可落地、安全可控的密钥保护方案，解决终端密钥管理难题。

[返回目录](#)

11. 数据库

11.1 多模数据库的前沿研究

随着互联网的不断发展，越来越多的物联网、电商、社交媒体等应用在丰富着人们的生活，这些应用每天持续的产生各种海量的异构数据(比如，表、图、JSON、XML、KV、Text 等)。为了满足这些海量异构数据的存储、利用、计算与分析，传统的做法是平台提供十几个不同的数据库产品分别满足相应需求，整体维护性和数据一致性管理成本可能因不断增加的数据量和数据类型而大大增加，进而影响到整个系统的使用。本课题旨在探索数据库平台的多样性数据存储管理优化问题。

建议研究方向：

- 1) 多模数据库的高效存储；
- 2) 多模数据库的元数据管理；
- 3) 多模连接查询和视图的最优化问题；
- 4) 基于工作负载的多模数据的统一查询、处理、计算和分析；
- 5) 基于 AI 的多模数据库。

[返回目录](#)

12. 网络技术

12.1 基于意图的下一代智能网络研究

本课题旨在探索 IBN(Intent-Based Network)理念在下一代智能网络管理系统中的实施，目标是基于大规模数据中心网络设备管理平台打造一个由意图解析、意图转译、意图实施、意图回馈等环节构成的闭环网络管理系统。

建议研究方向：



- 1) 大规模网络仿真业务方向：基于云计算环境构建数据中心网络模块级规模的仿真系统，在此仿真系统上开展除性能测试之外的绝大部分网络验证、故障分析、自动化测试工作；
- 2) 网络故障自愈 AI/ML 技术研究：研究如何根据历史网络故障（腾讯可提供历史故障集）构建网络故障样本集，基于此结构化样本集开展机器学习，不断积累经验，为新发故障提供快速、合理有效的解决措施。

[返回目录](#)