

一、国内研究现状与发展动态

国内在组合零样本识别领域的研究，整体以“技术实用化”为核心目标，依托成熟的人工智能技术基础，形成了快速响应实际需求、高效适配现有模型的研究特色。研究重点围绕如何将通用人工智能模型与具体识别任务结合，通过针对性优化提升技术在实际场景中的应用价值，构建了“基础模型适配+实用性能提升”的主流发展路径。

（一）核心研究现状

1. 跨模态信息融合能力持续提升：针对图像与文本两类信息的协同处理问题，国内研究通过优化语义关联挖掘与视觉特征提取技术，实现了图文信息的高效融合。相关研究通过强化对“属性-物体”关系的理解，有效提升了模型对未见过的组合概念的识别能力，相关成果已在多个主流应用场景中得到验证，形成了一批具有实用价值的技术方案。
2. 主流基础模型的适配研究深入：以CLIP等国际知名视觉语言模型为基础，国内研究团队重点开展了模型的本土化与场景化适配工作。通过优化模型的提示策略、特征表达形式及双模态协同机制，有效解决了通用模型在特定任务中的性能局限，形成了一系列兼顾效率与效果的改进方案，推动了基础模型在国内相关领域的落地应用。
3. 研究场景向实际需求贴近：国内研究已逐步摆脱传统“封闭测试环境”的局限，转向更符合真实应用的“开放场景”评估。当前研究不仅要求模型能识别已知组合，更强调在复杂多样的候选概念中精准定位未见过的组合，这种转变使研究成果更贴合智能检索、内容推荐等实际产业需求，提升了技术的落地价值。

（二）关键发展动态

国内研究正加速向“实用化突破”方向迈进，核心趋势体现在两个方面：一是充分利用大模型积累的海量知识，解决组合识别中存在的“常见组合依赖”“跨场景性能波动”等实际问题；二是强化技术的工程化适配能力，通过优化模型效率、降低部署成本，推动相关技术从实验室走向产业应用。整体而言，国内研究的技术落地速度与实用化程度正持续提升。

二、国外研究现状与发展动态

国外在组合零样本识别领域的研究以“理论创新”为核心驱动力，长期在基础研究、框架设计与前沿方向探索上处于引领地位。研究注重从本质上提升模型的泛化能力，形成了“底层机制突破—技术框架创新—评估体系升级”的完整研究链条，为全球该领域的发展提供了重要指引。

（一）核心研究现状

1. 基础理论与框架创新成果丰富：国外研究聚焦模型泛化能力的本质提升，提出了一系列全新的学习范式与理论模型。这些研究不再局限于对现有技术的局部改进，而是从模型的学习目标、网络结构、优化逻辑等底层维度进行重构，为解决组合零样本识别的核心难题提供了坚实的理论支撑。
2. 视觉语言模型的潜力深度挖掘：以CLIP为代表的视觉语言模型成为国外研究的核心工具，但与国内侧重适配不同，国外研究更聚焦挖掘模型的内在潜力。通过优化模型的提示交互方式、特征表达质量及多阶段处理逻辑，有效弥补了通用模型在细粒度识别、局部特征捕捉等方面的不足，进一步拓展了模型的应用边界。
3. 前沿评估体系逐步完善：国外率先提出“部分监督”“开放类别”等更具挑战性的评估标准，构建了覆盖简单到复杂场景的多元评估体系。通过系统分析模型在各类极端条件下的表现，明确了当前领域的核心技术瓶颈，为全球研究指明了攻坚方向，推动了整个领域向更高标准发展。

(二) 关键发展动态

国外研究正朝着两个核心方向推进：一是“多模态深度协同”，积极探索图像、文本、知识图谱等多类信息的融合机制，提升模型对复杂场景的综合理解能力；二是“极端泛化能力构建”，致力于让模型在极少先验信息的情况下，仍能高效识别全新的组合概念，推动模型向更贴近人类认知的智能水平发展。这些前沿探索为全球该领域的技术创新提供了重要参考。

三、研究总结与趋势展望

综合来看，国内外在组合零样本识别领域的研究呈现出鲜明的互补特征：国内以应用需求为导向，侧重技术的落地适配与实用效能；国外以理论突破为核心，聚焦基础机制的创新与前沿方向的引领。尽管研究侧重点不同，但两者的核心目标高度一致，均致力于突破模型在未知组合识别上的泛化能力瓶颈。

当前全球该领域已形成两大明确的共性趋势：一是视觉语言模型成为基础研究载体，各国研究团队均围绕其开展针对性优化；二是研究场景从封闭走向开放，从实验室模拟转向真实环境应用。未来，跨领域知识的融合应用、模型的轻量化与高效化、极端场景下的泛化能力提升等方向，将成为国内外共同关注的研究热点，这也为本次大创项目提供了清晰的创新方向与切入点，可依托国内外研究基础，聚焦实用化创新或理论应用突破，形成具有价值的研究成果。

三、研究现状总结与趋势展望

3. 关键发展动态

国外研究正朝着“多模态深度协同”与“极端泛化能力”方向推进，一方面探索视觉、语言、知识图谱等多源信息的融合机制，另一方面致力于让模型在“极少先验信息”的情况下实现高效识别，相关研究为全球领域发展提供了重要的技术参照。

三、国内外研究对比与共性趋势

国内外研究虽呈现“应用导向”与“理论导向”的差异化特征，但核心目标一致，均致力于提升模型在复杂场景下的泛化能力。当前领域已形成两大共性趋势：一是视觉语言模型成为基础工具，各国研究均围绕其进行针对性优化；二是“从封闭到开放、从实验室到真实场景”成为必然方向，研究越来越贴近实际应用需求。未来，跨领域知识融合、轻量化优化等方向将成为国内外共同的研究热点。