

报告正文

参照以下提纲撰写，要求内容翔实、清晰，层次分明，标题突出。
请勿删除或改动下述提纲标题及括号中的文字。

（一）立项依据与研究内容（建议 8000 字以下）：

1. 项目的立项依据（研究意义、国内外研究现状及发展动态分析，需结合科学研究发展趋势来论述科学意义；或结合国民经济和社会发展中迫切需要解决的关键科技问题来论述其应用前景。附主要参考文献目录）；

1.1 “需求牵引”

云南省与缅甸、老挝、越南三国接壤，边境线长达 4060 公里，边境地区山川秀丽，被誉为“植物王国”和“世界花园”。习近平总书记多次强调，“治国必治边”，边境安防工作的重要性贯穿于国家发展的各个阶段。然而，边境地区地形复杂多变，山地、高原、河谷等地貌交错分布且缺乏天然屏障，使得边境地区长期以来存在严重的非法越境隐患。非法越境问题主要表现为人口非法流动、毒品走私、货物走私以及跨境犯罪等多种形式，该问题不仅直接影响边境地区的安全与稳定，还可能引发跨境犯罪、社会矛盾和民族问题，进而威胁国家整体安全和社会秩序。近年来，云南边境地区非法越境问题愈发突出，给边境地区的社会稳定和国家的长治久安带来了巨大压力：

* **云南“1·04”组织外籍人员偷渡案。**在 2024 年 1 月，云南与江苏警方联合侦破一起组织毗邻国家人员经陆路偷渡入境的案件。**犯罪团伙利用中缅、中越边境地形复杂多变的密林和隐蔽通道**，组织外籍人员非法入境务工、通婚，同时衍生走私、涉毒等犯罪，查获走私物品案值 310 万元，缴获毒品 798 克，破坏边境社会稳定与经济安全^①（参见图 1(a)-(b)）。

* **打洛边境派出所的防疫战。**2017 年起，打洛镇开始实施立体化管控，在陆地上修筑栅栏、建设抵边警务室，边境栅栏长达 5 公里，陆地上的偷渡渠道逐渐被封锁，越境路径全被挤压到了水路，因此河面设有雷达报警。在此基础上，民警定期轮岗和巡逻，**打造了立体式的监控网络**，成为云南边防的典范。然而近年仍有非法越境的情况发生，例如，疫情期间因两例

^①信息来源：江西公安公众号 (<https://news.qq.com/rain/a/20250214A01KFX00>)

缅甸籍确诊病例非法入境，边境小城瑞丽不得不封城一周。从云南边境整体情况来看，2024年上半年云南边检机关共侦破妨害国（边）境管理刑事案件 893 起，抓获犯罪嫌疑人 2930 人，打掉组织偷渡团伙 262 个，非法越境活动在云南边境地区仍然存在一定的规模^②（参见图 1(c)-(d)）。

*** 云南“1·24”组织运送他人偷越国（边）境重特大案件。**2020 年至 2021 年期间，张某、匡某等 11 人不顾疫情防控政策，为获取高额报酬，违反出入国（边）境管理法规，针对监控设施覆盖的漏洞和盲区，采取带路爬山、以摩托和轿车交互运输、用货物遮挡、绕道小路等方式逃避检查，组织、运送大批人员非法出、入境，更有涉案人员趁机实施运输毒品犯罪，严重破坏国（边）境管理和疫情防控工作^③（参见图 1(e)-(f)）。

上述案件类型显示了执法部门在追捕犯罪嫌疑人和保护受害者方面的积极努力，但也反映出云南边境地区在打击非法越境问题方面仍存在提升空间。



图 1. 云南边防现状

习近平总书记在中共中央政治局第十八次集体学习时强调，“推进边疆治理体系和治理能力现代化，是中国式现代化的应有之义”。打击非法越境是一个超大场景下全天候的目标发现任务，监控系统需要在几十到几千米、甚至几十千米的较大地理空间范围内对目标进行监控，还需要实现 7x24 小时全气候条件下的监控覆盖。边境公安针对该问题正积极探索并应用前沿监控技术，以替代传统的警员巡逻模式，旨在克服传统手段中巡逻覆盖范围有限、反应速度滞后以及难以实现全天候、全方位监控等固有缺陷，同时有效应对日益隐蔽化和智能化的非法过境手段。然而，现有监控设备在技术层面仍存在显著不足，特别是缺乏高科技手段（如智能识别系统、多光谱成像、无人机协同巡查等）

^②信息来源：红星新闻 (<https://news.qq.com/rain/a/20201005A08HYK00>)

^③信息来源：云南长安网 (https://www.yncaw.gov.cn/html/2022/ftnw_0226/87066.html)

的深度集成，导致对非法过境行为的精准判断与打击能力未能达到预期效果。因此，目前超大场景安防监控仍有短板，并呈现如下现实需求：

*** 需求 1：适应超大场景复杂地形和动态环境变化的监控需求。**从云南“1·04”组织外籍人员偷渡案可知，由于边境地区地形复杂多样，涵盖高山、丛林、河流等多种地貌，且环境动态变化频繁，导致监控设备在应对超大场景内物体随天气、时间变化导致的光线、能见度以及物体形态变化时，监控算法适应性和稳定性急剧下降或失效。这不仅限制了设备对环境的感知能力，还影响了监控系统的稳定性和可靠性，进而制约了边境监控的效能。因此，边境安防迫切需要开发能够适应复杂地形和动态环境变化的监控设备，以提升边境监控系统的整体性能和可靠性；

*** 需求 2：超大场景下对潜在目标的实时预测和有效预判的监控需求。**从打洛边境派出所的防疫战案例可知，现有监控设备在超大场景内的潜在目标主动预测与定位能力存在显著不足。当前监控系统主要依赖被动式监控模式，缺乏基于数据分析的主动预测机制，难以对非法越境目标的出现位置、行为轨迹及潜在风险进行实时预测与精准预判。这一技术短板导致边境安防工作长期处于被动监控模式，难以在非法活动发生前采取有效干预措施，从而降低了整体防控效能。因此，边境安防迫切需要研发具备主动预测与定位能力的监控系统，通过大数据分析和智能化技术，实现对潜在目标的实时预测和有效预判，从而提升边境安防的主动防御能力和整体防控效能；

*** 需求 3：超大场景下主动搜索和智能调控的监控需求。**从云南“1·24”组织运送他人偷越国（边）境重特大案件可知，现有监控设备体系存在明显的局限性。一方面，监控范围狭小导致在超大场景下的监控部署与运维成本居高不下，尤其是在云南省 4060 公里长的边境线上，单一摄像头的监控范围受限，加之山地、树林等物体的遮挡，难以实现无缝覆盖。另一方面，现有设备主要依赖被动式监控模式，无法通过调整摄像头的俯仰角、转动角及焦距主动搜索和发现超大场景下的目标，仍需依赖人力调整。因此，亟需研发具备宽广监控范围、主动搜索能力和智能化调控功能的监控系统，以降低监控部署与运维成本，进一步提升边境地区的整体防控效能。

综上，本项目旨在通过运用大数据分析与机器学习技术，将人工干预的被动监控模式转变为算法实时控制的主动预测与搜索模式，提升监控设备在超大场景下的自主监控能力，努力满足上述三个方面的现实边防监控需求，从而有效治理非法越境问题，保障国家主权安全，推动边疆地区经济社会发展，为国家的稳定和发展创造良好环境。

1.2 “突破瓶颈”

本项目以实际需求为导向，充分考虑边境安防系统的前述需求，在硬件设备基础上设计软件算法，优化和改善现有方案的不足。超大场景下的主动目标预测与搜索核心步骤如下：（1）利用边防监控摄像设备，通过调整摄像云台位置（如果定点监控则位置固定不变），焦距、偏转角和俯仰角，得到超大场景内子区域图像，分析并提取图像特征；（2）利用图像特征预测潜在目标相对当前超大场景内子区域的位置（3）规划路径，通过调整摄像云台使得摄像设备对准目标，完成目标在超大场景内的搜索。下面从硬件和软件两个方面阐述现有方案的不足与困难。

监控摄像设备硬件行业近年来经历了显著的发展。监控摄像机已经具备了更高的分辨率、动态范围和低光性能。更重要的是，智能化功能如高清图像处理、人脸识别、智能分析等技术的不断发展，使得视频监控系统的功能愈发强大。通过深度学习、计算机视觉等技术的融合应用，视频监控设备具备了更精准的目标识别、行为分析和异常检测能力。

然而，云南边境地区地形复杂，植被覆盖丰富，对监控设备的部署和使用提出了特殊要求。云南省地势西北高、东南低，自北向南呈阶梯状逐级下降，地形地貌复杂多样，包括滇西北的高海拔地区、滇西南的低海拔热带雨林地区等。这些地区植被茂密，地形崎岖，传统的重型监控设备在部署和维护上存在诸多困难。因此，轻量化、便携式的监控设备在云南边境地区具有迫切的需求。轻量化设备不仅便于携带和安装，还能适应复杂的地形和植被环境，提高监控效率和准确性。

针对监控硬件设备，本项目针对云南边境地区的特点，拟采用监控范围广，便携轻量的摄像设备，灵活的开展面向超大场景的边境安防任务，如图 2（a）所示。

超长变焦能力的监控设备可以在不改变监控位置的情况下，通过调整焦距来捕捉远距离的清晰画面，为超大场景安防监控提供了必要的前提。例如，现

有主流便携式长变焦摄像机已实现 24-3000 mm 等效变焦能力，在能见度较好的天气能够发现 10-15 千米外的人形目标，如图 2（b）-（c）所示。这种超长变焦能力使得监控设备能够在复杂的地形和植被环境中，有效监控远距离的目标，提高监控的效率和准确性。同时，轻量化和便携式的监控设备也便于集成到无人机或移动机器人上。这种集成方式可以在复杂的遮挡场景中提供更加灵活的监控解决方案。例如，无人机机载摄像设备能够快速调整监控位置、角度和焦距，适应不同的监控需求，这种灵活性使得监控设备能够在复杂的环境中实现高效监控，提高安防系统的整体性能。该硬件方案与目前主流的远距离监控摄像设备相比，在成本、体积和重量等方面具有明显优势，更适合云南边境超大场景安防任务，见表 1。

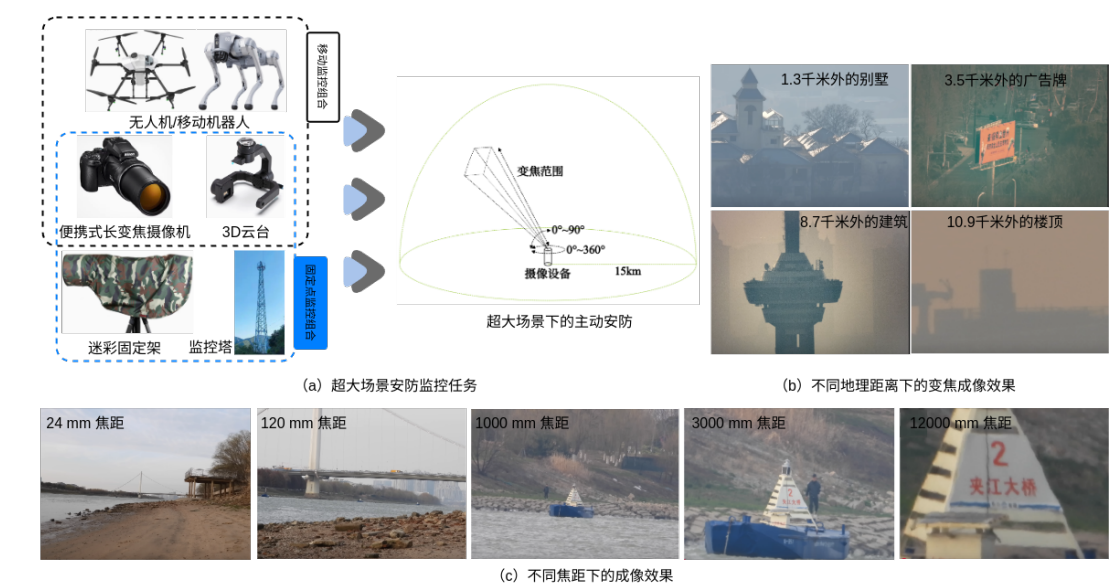


图 2. 超大场景边防任务

表 1. 摄像设备对比

设备	价格（人民币）	重量	监控范围	安装类型	故障售后
远距离监控摄像机	47.7 万（5 台起订）	110 KG	15KM	地面	研发机构
便携式长变焦摄像机	1.05 万（量产机）	1.9 KG	15KM	地面/机载	全球门店

现有监控软件及智能化程序方面，主要围绕以下三个相关领域展开：（1）超大场景下变化物体的图像特征提取（2）超大场景下的动态目标预测（3）面向超大场景的目标主动搜索。然而，在大量文献调查后，申请人发现变化物体有效特征提取、动态环境下的目标主动预测与搜索相关研究较少，同时上述图像特征提取、目标主动预测与搜索 3 个领域的相关技术瓶颈也给本课题带来了困难与挑战。具体分析如下：

(1) 挑战一：现有图像特征提取方法在超大场景下对变化物体特征提取鲁棒性和泛化能力较差，往往对变化物体和目标识别不完全，给后续潜在目标的预测和搜索带来挑战。

近年来，基于 Transformer 的视觉模型 [1-9] 通过自注意力机制显著提升了图像特征提取的全局建模能力，但其对动态环境下的变化物体 (位置和形态等的变化) 特征提取能力因分布变化导致鲁棒性和泛化能力大幅下降。同时，对边境安防实时处理和相应的需求，轻量化模型 (如 DDRNets[10]) 通过降低计算复杂度实现了实时处理，但其在动态环境中的鲁棒性受限于特征表达的局部性。图 3 所示为已训练的 DDRNets 模型 [10] 对新场景的分割效果，DDRNets 对新场景下的大部分物体都没有完全识别并分割，只有物体的部分区域分割正确。



图 3. 已训练的 DDRNets 模型对新场景的分割效果

图像特征提取的其鲁棒性和泛化能力在目标预测和搜索中至关重要，因为它们能够有效突出关键信息，确保预测和搜索算法在复杂多变的边境环境中稳定运行。为应对这一问题，现有方法主要通过数据增强，模型优化，迁移学习几种途径实现。(1) 数据增强方面，对原始图像进行数据合成 (如随机图像变换 [11]、噪声添加 [12] 等) 或通过学习的方式 (如 CLUT-Net[13]) 增加数据的多样性和数量，进而提升模型的鲁棒性和泛化能力。(2) 模型优化方面，SGNet [14]、CPP-Net [15]、DynamicCity[16]、CTA-Net[17] 等新方法通过引入独特的网络结构设计、多尺度特征融合以及注意力机制等，显著提升了图像特征提取模型的泛化能力。(3) 迁移学习方面，CLIP-Adapter[18]、TOAST[19] 等方法采用轻量级策略，高效地将预训练模型的知识迁移到新的任务或领域中，极大地增强了模型的泛化能力和适应性。此外，组合式特征提取方法优化也有新

进展, CSFF[20]、DilateFormer[21] 等通过多尺度特征提取以及结合传统与深度学习方法, 在增强特征的鲁棒性方面具有重要意义。同时, 像 FFCA-YOLO [22] 结合了特征增强模块、特征融合模块和空间上下文感知模块, 有效增强了小目标的特征表达并抑制复杂背景干扰。这些新方法从不同角度为提升图像特征提取的鲁棒性和泛化能力做出了贡献。

然而, 在处理图像损坏、噪声干扰、小样本量以及跨领域任务时, 现有方法仍难以充分捕捉因果特征, 导致在未见域中的表现欠佳。这些挑战表明, 进一步优化特征提取方法以增强模型的泛化能力仍是未来研究的重要方向。本项目拟从特征抽象的视角出发, 结合图像中物体之间的语义关系与物体局部图像特征, 利用图结构表示并充分捕捉变化物体的不变特征, 提升模型对复杂多变场景的泛化能力。

(2) 挑战二: 现有动态图学习方法在超大场景下对稀疏的观测信息利用不够充分, 给目标潜在位置预测精度和可解释性带来挑战。

动态图神经网络 (Dynamic GNN) 以及动态图学习领域, 主流的特征提取方法主要使用多尺度动态特征提取方法。例如, MS-STGNN [23] 基于多尺度时间特征提取的图神经网络, 通过结合多尺度的时间特征来增强模型对动态图的建模能力。另外, 动态图学习效率方面, WinGNN[24] 通过引入随机梯度聚合窗口, 提升了动态图神经网络在时间序列建模中的效率和准确性; SEIGN 给出面向大规模动态图学习的高效动态图学习方案。

与超大场景目标预测相关动态图学习的研究也取得了一定进展, 例如在轨迹预测方面, 基于时空区间信息的 Seq2Seq[25] 框架试图解决轨迹恢复问题, 通过结合时空特征和图结构信息, 提升了轨迹预测的精度和鲁棒性。还有研究关注即时轨迹预测, 如 ITPNet[26], 通过集成 HiVT[27] 和 LaneGCN[28] 等骨干网络模型, 提升了轨迹预测的准确性。

然而, 这些方法在超大场景下的稀疏观测条件下, 目标潜在位置预测精度不高且缺乏可解释性。如果预测结果缺乏可解释性, 那么将使得后续超大场景下的主动搜索过度依赖数据本身, 一旦数据中存在噪声和多样环境下的不确定因素, 将极大限制设备主动搜索目标的精度。动态图的可解释相关的研究旨在清晰地理解、解释动态图模型从输入数据中如何学习模式、获取重要模式、以及模型做出预测的过程, 利用稀疏注意力机制、因果模型 [29]、对比学习 [30] 和缓冲区更新技术, 描述数据样本贡献性 [31]、特征重要性 (如 DyExplainer[30])、

模型结构重要性（如 E-SAGE[32]）、时间维度的影响（如 T-GAP[33]）、以及应用领域的特殊需求。

面对超大场景下的安防任务，只有提供可解释的可信预测，才能为目标主动搜索提供提供依据。因此，本项目拟使用 GNN 对动态图结构利用迭代搜索的方式在稀疏的数据样本下获得有效的特征表示，并由此预测目标位置的分布，在此基础上加入 GNN 可解释机制，得到可信的目标位置预测。

（3）挑战三：现有方法对多条件约束下无法实时处理和规划最优路径，给监控设备主动搜索潜在目标带来挑战。

主动目标搜索利用之前对潜在目标位置的预测，规划最优路径并控制摄像设备指向目标。经典路径规划方法在动态多目标场景中面临严重局限，在高频重规划情境下算法的计算开销呈指数级增长 [34]、对密集动态目标时容易产生路径震荡问题 [35]、或求解时间不可控 [36]，难以满足实时需求。

近年来，数据驱动的方法在路径规划领域取得显著进展（如 Neural A* [37], Value Iteration Networks[38]）。基于深度学习和强化学习的路径优化算法也逐渐成为研究热点 [39–41]。基于深度学习的主动搜索方法 [42]（Active Search with Deep Learning, ASDL）通过结合强化学习与卷积神经网络（CNN），显著提升了目标搜索的精度与效率。PriorSparseR-CNN 框架 [43] 通过区域生成网络（Region Generation Network, RGN）和特征聚合策略，优化了稀疏目标检测的性能，但其计算复杂度较高，难以满足实时性需求。同时，多目标爆搜方法（Multi-Objective Explosive Search, MOES）[44] 通过生成多个候选目标并结合目标跟踪算法，实现了高精度检测，但其存储空间消耗大，且在处理动态场景时易受目标形变与遮挡影响。在考虑约束条件的路径规划方面，现有研究能够处理单一或多约束条件下的优化，如自动驾驶场景下在安全性、效率和舒适性之间找到平衡 [45]。社会感知运动规划 [46]（Socially Aware Motion Planning）框架融合深度强化学习技术，使机器人能够理解和预测人类行为，将人的行为作为约束，生成满足条件的导航轨迹，但计算复杂度较高。随后，基于视觉输入的城市导航策略 [47] 实现了无地图导航方法，通过强化学习直接从视觉输入学习导航策略，适用于复杂的城市环境，但对动态环境下的物体处理能力有限。Think2Drive[48] 提出了在潜在世界模型中进行”思考”的强化学习框架，通过建立环境的压缩表示，显著减少了对大量物理仿真的依赖，提高了学习效率。

尽管路径优化问题的研究取得了诸多进展，但仍存在以下不足：首先，现

有算法在处理大规模动态路径优化问题时，往往面临计算复杂度高、收敛速度慢的问题，难以满足实时性要求。其次，路径优化模型的普适性和灵活性不足，现有方法多针对特定场景设计，难以适应多样化的实际需求。再次，路径优化中的不确定性因素（如交通状况、天气变化）尚未得到充分解决，现有模型的鲁棒性和适应性有待进一步提升。最后，路径优化技术的实际应用仍面临数据安全和隐私保护等挑战，如何在保证数据安全的前提下实现高效优化，仍需进一步研究。因此，本项目拟利用强化学习强大的规划与决策能力，设计适用于多条件约束下的快速训练方法，以支撑边境超大场景下的实时控制和搜索。

参考文献

- [1] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Houtsby N. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [2] Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S. End-to-End Object Detection with Transformers[J]. arXiv preprint arXiv:2005.12872, 2020.
- [3] Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows[J]. arXiv preprint arXiv:2103.14030, 2021.
- [4] Zhu X, Su W, Lu L, Li B, Wang X, Dai J. Deformable DETR: Deformable Transformers for End-to-End Object Detection[J]. arXiv preprint arXiv:2010.04159, 2021.
- [5] Zheng S, Lu J, Zhao H, Zhu X, Luo Z, Wang Y, Fu Y, Feng J, Xiang T, Torr P H, others. Rethinking Semantic Segmentation from a Sequence-to-Sequence Perspective with Transformers[J]. arXiv preprint arXiv:2012.15840, 2021.
- [6] Bao H, Dong L, Piao S, Wei F. BEiT: BERT Pretraining of Image Transformers[J]. arXiv preprint arXiv:2106.08254, 2021.
- [7] He K, Chen X, Xie S, Li Y, Dollár P, Girshick R. Masked Autoencoders Are Scalable Vision Learners[J]. arXiv preprint arXiv:2111.06377, 2021.
- [8] Liu Y, Zhang Y, Wang Y, Hou F, Yuan J, Tian J, Zhang Y, Shi Z, Ji X, Wang J. Tokens-to-Token ViT: Training Vision Transformers from Scratch on ImageNet[J]. arXiv preprint arXiv:2101.11986, 2021.
- [9] Wang W, Xie E, Li X, Fan D-P, Song K, Liang D, Lu T, Luo P, Shao L. Pyramid Vision Transformer: A Versatile Backbone for Dense Prediction without Convolutions[J]. arXiv preprint arXiv:2102.12122, 2021.
- [10] Pan H, Hong Y, Sun W, Jia Y. Deep Dual-Resolution Networks for Real-Time and Accurate Semantic Segmentation of Traffic Scenes[J]. IEEE Trans. Intell. Transp. Syst., 24(3): 3448–3460, 2023.
- [11] Cubuk E D, Zoph B, Shlens J, Le Q V. RandAugment: Practical automated data augmentation with a reduced search space[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020: 702–703.

- [12] Chowdhury S, Soni B. R-VQA: A robust visual question answering model[J]. Knowledge-Based Systems, 309 : 112827, 2025.
- [13] Mei X-B, Liu X, Chang X, Yang X, Tang Z-H. CLUT-Net: Learning adaptively compressed representations of 3DLUTs for lightweight image enhancement[C]. Proceedings of the ACM International Conference on Multimedia. 2022 : 2914–2922.
- [14] Liu D, Cui Y, Tan W, Chen Y. SG-Net: Spatial granularity network for one-stage video instance segmentation[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021 : 9816–9825.
- [15] Guo Z, Gan H. CPP-Net: Embracing multi-scale feature fusion into deep unfolding CP-PPA network for compressive sensing[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2024 : 25086–25095.
- [16] Bian H, Kong L, Xie H, Liu Z, Qiao Y. DynamicCity: 4D Scene Generation via Spatiotemporal Feature Compression[J]. ICLR, 2025.
- [17] Meng C, et al.. CTA-Net: A CNN-Transformer Aggregation Network for Improving Multi-Scale Feature Extraction[J]. arXiv preprint arXiv:2410.11428, 2024.
- [18] Gao P, Zhang H, Lu S, Liu Z. CLIP-Adapter: Better vision-language models with feature adapters[J]. arXiv preprint arXiv:2110.04544, 2021.
- [19] Shi B, Gai S, Darrell T, Wang X. TOAST: Transfer Learning via Attention Steering[J]. arXiv preprint arXiv:2305.15542, 2023.
- [20] Cheng G, Yao X, Guo L. Cross-Scale Feature Fusion for Object Detection in Optical Remote Sensing Images[J]. IEEE Geoscience and Remote Sensing Letters, 17(3) : 431–435, 2020.
- [21] Jiao J, others. DilateFormer: Multi-Scale Dilated Transformer for Visual Recognition[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023 : 8906–8919.
- [22] Yin Z, Mu Y, Guiyi Z, Yong L, Pengyu G, Junhua Y. FFCA-YOLO for Small Object Detection in Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 62(2) : 1234–1245, 2024.
- [23] Wang Y, Kong Y, Tang X, Chen X, Xu Y, Chen J, Sun S, Guo Y, Chen Y. A Multi-Scale Spatial-Temporal Graph Neural Network-Based Method of Multienergy Load Forecasting in Integrated Energy System[J]. IEEE Transactions on Industrial Informatics : 2652–2666, 2023.
- [24] ZHU Y, CONG F, ZHANG D, GONG W, LIN Q, FENG W, DONG Y, TANG J. WinGNN: Dynamic Graph Neural Networks with Random Gradient Aggregation Window[C/OL]. Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23). 2023 : 13 pages. doi:10.1145/3580305.3599551.
- [25] Anonymous. A Graph-based Representation Framework for Trajectory Recovery via Spatiotemporal Interval-Informed Seq2Seq[J]. arXiv preprint arXiv:2410.15666, 2024.
- [26] Anonymous. ITPNet: TOWARDS INSTANTANEOUS TRAJECTORY PREDICTION FOR AUTONOMOUS DRIVING[J]. arXiv preprint arXiv:2410.15666, 2024.

- [27] Zhou Z, Ye L, Wang J, Wu K, Lu K. HiVT: Hierarchical Vector Transformer for Multi-Agent Motion Prediction[C]. CVPR. 2022.
- [28] Liang M, Yang B, Hu R, Chen Y, Liao R, Feng S, Urtasun R. Learning lane graph representations for motion forecasting[C]. ECCV. 2020.
- [29] Zhao K, Zhang L, Liu Z. Causality-Inspired Spatial-Temporal Explanations for Dynamic Graph Neural Networks[C]. ICLR. 2024.
- [30] Wang T, Luo D, Cheng W, Chen H, Zhang X. DyExplainer: Explainable Dynamic Graph Neural Networks[C]. WSDM. 2024.
- [31] Chen Z, Li P, Liu H, Hong P. Characterizing the Influence of Graph Elements[C]. The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023. [S.l.] : OpenReview.net, 2023.
- [32] Wang T, Luo D, Cheng W, Chen H, Zhang X. E-SAGE: Explainability-based Defense Against Backdoor Attacks on Graph Neural Networks[C]. WSDM. 2024.
- [33] Jung J, Jung J, Kang U. Learning to Walk across Time for Interpretable Temporal Knowledge Graph Completion[C]. KDD. 2021.
- [34] Bhardwaj M, Choudhury S, Scherer S. Learning Heuristic Search via Imitation[Z]. 2017.
- [35] Qureshi A H, Miao Y, Simeonov A, Yip M C. Motion Planning Networks: Bridging the Gap Between Learning-Based and Classical Motion Planners[J]. IEEE Transactions on Robotics, 37(1): 48 – 66, 2021.
- [36] Li T-M, Aittala M, Durand F, Lehtinen J. Differentiable Monte Carlo ray tracing through edge sampling[J]. ACM Trans. Graph., 37(6), 2018.
- [37] Archetti A, Cannici M, Matteucci M. Neural Weighted A*: Learning Graph Costs and Heuristics with Differentiable Anytime A*[M/OL]. Machine Learning, Optimization, and Data Science. [S.l.] : Springer International Publishing, 2022 : 596 – 610. http://dx.doi.org/10.1007/978-3-030-95467-3_43.
- [38] Tamar A, Wu Y, Thomas G, Levine S, Abbeel P. Value Iteration Networks[Z]. 2017.
- [39] Zheng J, Ding M, Sun L, Liu H. Distributed stochastic algorithm based on enhanced genetic algorithm for path planning of multi-UAV cooperative area search[J]. IEEE Trans. Intell. Transp. Syst., 24 : 8290 – 8303, 2023.
- [40] Cui J, Wu L, Huang X, Xu D, Liu C, Xiao W. Multi-strategy adaptable ant colony optimization algorithm and its application in robot path planning[J]. Knowl.-Based Syst., 288 : 111459, 2024.
- [41] Fang W, Liao Z, Bai Y. Improved ACO algorithm fused with improved Q-Learning algorithm for Bessel curve global path planning of search and rescue robots[J]. Robot. Auton. Syst., 182 : 104822, 2024.
- [42] Hottung A, Kwon Y, Tierney K. Efficient Active Search for Combinatorial Optimization Problems[C]. ICLR. 2022.
- [43] Wang Q, Qian Y, others. Apply prior feature integration to sparse object detectors[J]. Pattern Recognition, 2024.

- [44] Unknown. Multi-Objective Explosive Search: A Novel Approach for High-Precision Target Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
- [45] Yang R, Sun X, Narasimhan K. A Generalized Algorithm for Multi-Objective Reinforcement Learning and Policy Adaptation[Z]. 2019.
- [46] Chen Y F, Everett M, Liu M, How J P. Socially aware motion planning with deep reinforcement learning[C]. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2017, Vancouver, BC, Canada, September 24-28, 2017. [S.l.] : IEEE, 2017 : 1343 – 1350.
- [47] Mirowski P, Grimes M K, Malinowski M, Hermann K M, Anderson K, Teplyashin D, Simonyan K, Kavukcuoglu K, Zisserman A, Hadsell R. Learning to Navigate in Cities Without a Map[Z]. 2019.
- [48] Li Q, Jia X, Wang S, Yan J. Think2Drive: Efficient Reinforcement Learning by Thinking in Latent World Model for Quasi-Realistic Autonomous Driving (in CARLA-v2)[Z]. 2024.

2. 项目的研究内容、研究目标，以及拟解决的关键科学问题（此部分为重点阐述内容）；

本项目紧密围绕实际需求开展研究，深入剖析当前已有方法在应对现实挑战时的局限性，旨在提出创新性研究方案，以攻克现存难题，总体研究思路如图 4所示，拟由此提升边防监控的智能化水平。

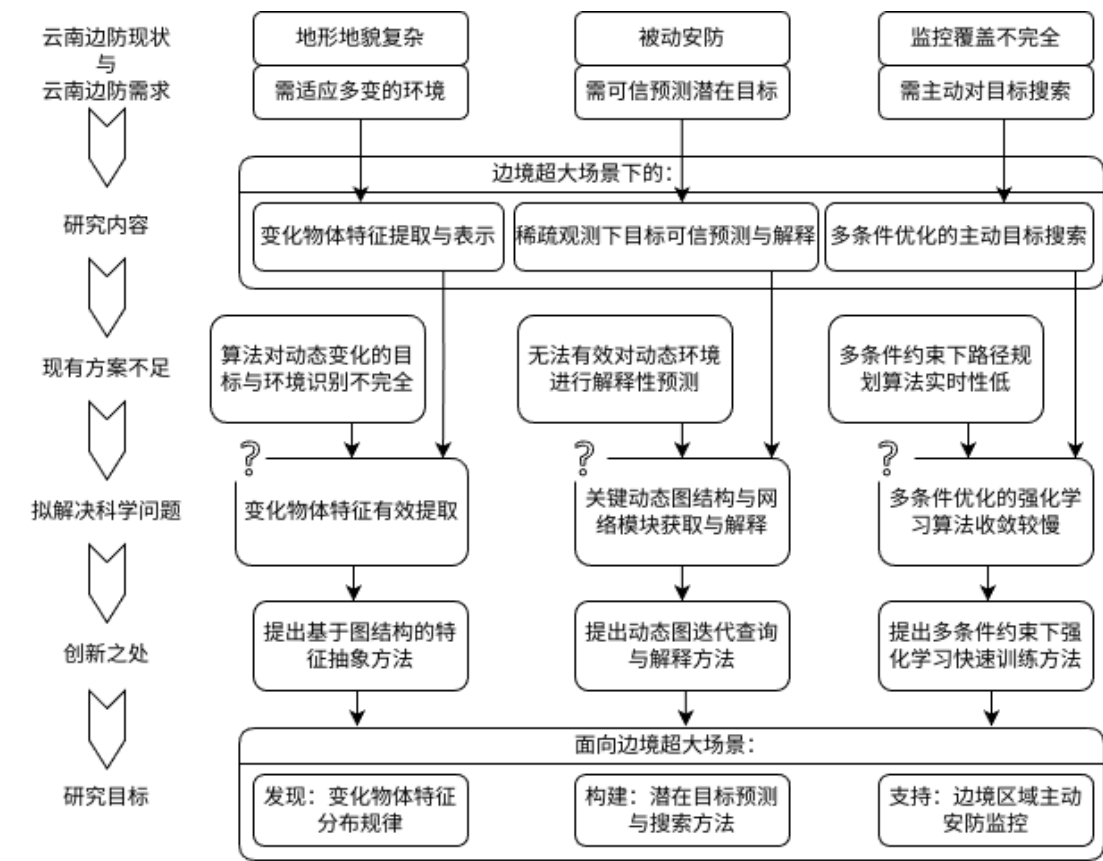


图 4. 组织结构

2.1 研究内容

(1) 超大场景下变化物体特征提取与表示方法研究

物体特征的有效提取与表示是后续目标预测与主动搜索的前提，其目的是为后续过程提供不受环境因素和场景变化影响的物体特征表示。在超大场景下，同一物体的特征往往那个会受到光照，天气，周围物体遮挡或物体本身的形态变化而产生较大变化，使得该物体不能有效被视觉模型识别和理解。因此，研究物体特征抽象与分离方法，使复杂变化对物体特征的影响最小化。一方面，将物体内在的固有特征和物体受到外界因素变化显现的外在特征相分离，另一方面，使用对比学习等手段进一步同化相同物体的表示，异化不同物体的表示，以实现对变化物体的成功识别和理解。由此解决现有方法对超大场景下变化物

体特征提取的困难，为后续目标的预测和主动搜索提供稳定可靠的物体表示。

(2) 超大场景稀疏观测下的目标位置可信预测与解释方法研究

目标位置可信预测根据当前摄像设备的局部观测信息推断目标物体在超大场景内位置的概率分布。摄像设备的局部观测往往具有稀疏性，因为单一监控设备往往不能全面清晰的获取超大场景任意时刻的全景，仅能够观测并获取其中的一个子区域的信息。目标可信预测的目的是为后续目标主动搜索提供必要依据，否则搜索过程将冗余和低效，降低目标发现的可能性。预测的目标位置依赖于当前观测的信息，但由于数据噪声和环境物体的不确定性，需要必要的可解释机制描述目标潜在位置的可信度。因此，研究观测与物体潜在位置之间的条件依赖关系，针对超大场景下的观测的稀疏性问题，重点研究当前观测物体对潜在目标之间的条件分布变化规律，探索并发现依赖关系背后的可解释性原理，为全数据流程各个环节的构建提供依据。

(3) 超大场景下多条件约束的主动目标搜索方法研究

主动目标搜索的目的有两个方面。首先，利用自动智能算法代替现有被动人工介入的监控模式，实现全天候监控覆盖；其次，利用单一超大场景主动搜索机制代替同能监控能力下大量被动监控设施。从而双管齐下的节约成本并提升监控能力。现有主动搜索方法主要采用区域覆盖的方式排查目标，对于超大场景下的动态目标搜索能力有限。因此，研究超大场景下动态目标的主动搜索，针对监控设备在调整转动角，俯仰角及变焦时的物理噪声和精度问题，重点研究自适应的控制方法，依据潜在目标的位置，通过序列化决策主动搜索并发现目标，解决现有方法存在的困难。

2.2 研究目标

本课题拟通过构建①超大场景下变化物体特征提取与表示方法；②超大场景下基于稀疏观测的目标位置预测与可解释方法；③超大场景下主动目标搜索方法方法，拟达到以下目标：(1) 发现变化物体特征之中的通用特征，进而深度模型对动态超大场景内物体的范化能力和稳定性；(2) 提出观测条件受限情况下的特定目标在超大场景中的潜在位置可信预测与主动搜索方法，将当下边境安防监控的被动监视状态转变为主动监控状态，提前对潜在目标进行响应；(3) 搭建实验平台、构建公开数据集、开发原型系统以提升研究深度广度、促进学术交流与技术应用落地。

2.3 拟解决的关键科学问题

上述研究内容在解决现有方法面对的困难和挑战时，主要需解决如下关键科学问题。

(1) 超大场景下变化物体图像特征提取有效性受环境变化显著下降的问题

在图像特征提取过程中，现有基于 CNN 的方法通过多层卷积操作提取物体特征，每一层卷积核能够捕捉图像中的局部特征（如边缘、纹理等），并通过堆叠多层网络逐步提取更高层次的语义信息（如形状、结构等）。然而在不断变化的超大场景下，物体的轮廓，边缘等关键信息往往会发生显著变或表现为较强的环境噪声。CNN 对噪声敏感，原因在于其卷积操作本质上是对局部像素的加权求和，噪声会直接干扰局部像素值，导致特征提取的偏差。同时，噪声可能在高层次特征中被放大，进一步影响模型的分类或检测性能。因此，噪声的存在会降低 CNN 特征提取的鲁棒性和准确性。基于 Transformer 的视觉模型虽然考虑到了细微像素点特征长距离依赖关系，但仍然无法有效处理较大的物体变化，对环境噪声依然较为敏感，原因与 CNN 模型类似。因此，为了避免单一局部特征受到噪声影响导致的模型误判，本项目拟对物体特征进行抽象，提取和利用不用变化下统一物体固有的不变特征，由此便能够准确识别物体。同时，辅以物体随环境变化的特征，便能够在正确识别物体的前提下，充分考虑原始局部特征，为后续预测提供更加全面细致的特征。

(2) 超大场景下动态特征学习过程中关键图结构与模型网络模块的获取与解释问题

超大场景通常具有数据规模庞大、环境复杂多变、目标动态性强等特点，这对特征学习的实时性、鲁棒性和可解释性提出了极高的要求。首先，超大场景中的数据流通常具有高维度和高复杂性的特点，传统特征学习方法难以有效捕捉数据中的动态变化。例如，在边境安防过程中，植物、动物、道路状况等目标的动态变化需要实时捕捉和分析，这对特征学习的实时性和鲁棒性提出了更高要求。其次，动态特征学习的可解释性问题尚未得到充分解决。深度学习“黑箱”特性使得特征提取过程缺乏透明性，难以解释模型决策的依据，这在边境安防监控等安全敏感领域中尤为关键。本项目拟提出一种基于动态图神经网络（Dynamic GNN）和可解释性人工智能（XAI）技术的动态特征学习方法。首先，利用动态图神经网络对超大场景中的数据流进行建模，通过图结构的动态更新捕捉目标之间的时空关系，提升特征学习的鲁棒性和适应性。其次，结

合可解释性人工智能技术，对特征提取过程进行可视化分析，揭示模型决策的关键因素，增强模型的可解释性和可信度。此外，通过设计轻量化的特征学习框架，降低计算复杂度，满足超大场景下的实时性需求。最后，通过实验验证该方法在复杂场景下的性能，并与现有方法进行对比，展示其在特征学习精度和可解释性上的优势。

(3) 超大场景与多条件约束下的强化学习算法收敛性问题

超大场景下多条件约束的主动搜索问题是智能监控等领域的核心挑战之一，其目标是在复杂环境中高效搜索和发现目标。传统搜索方法在超大场景下面临搜索空间大、目标分布稀疏以及多条件约束的挑战，导致效率低下。大空间搜索往往需要遍历的区域广阔，计算复杂度高；目标分布稀疏则增加了定位目标的难度，容易产生大量无效搜索。此外，多条件约束（如时间、资源限制和环境动态变化）进一步限制了搜索的灵活性和适应性，使得传统方法难以在复杂场景中高效完成任务。因此，本项目拟使用基于强化学习的主动搜索方法，通过智能体间的信息共享与协作，动态调整搜索策略以应对多条件约束。针对训练过程，需要设计合理的训练方法帮助模型在有限的资源下收敛。为了更好的适应物理监控设备的操控特性，研究将进一步优化多条件约束下的搜索策略，推动该技术在边境安防监控领域的应用。

3. 拟采取的研究方案及可行性分析（包括研究方法、技术路线、实验手段、关键技术等说明）；

3.1 研究方案与技术路线

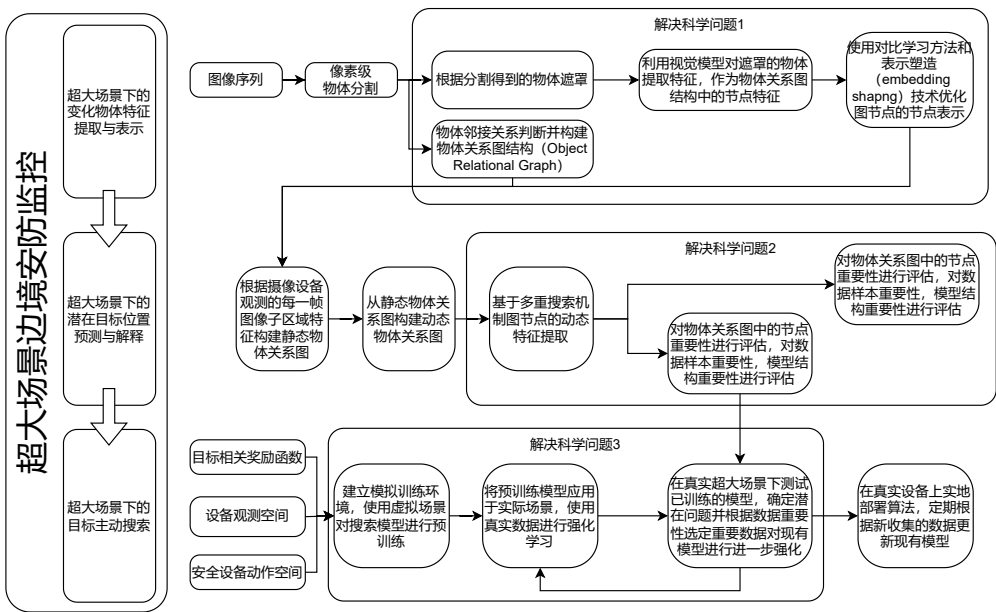


图 5. 总体框架图

3.2 研究方法

3.3 可行性分析

4. 本项目的特色与创新之处；

(1) 特色

* **紧迫性：**随着全球安全形势的日益复杂，跨境犯罪、恐怖主义等活动对国家边防安全构成了严重威胁。研发先进的边防主动预测与搜索技术，可以有效提升边防部门对这些威胁的感知和应对能力，保障国家的安全和稳定。因此，加快边防主动预测与搜索技术的研发，对于提升国家边防安全水平具有重要意义。

* **推广应用前景：**边防主动预测与搜索技术在未来的推广应用前景非常广阔。首先，边防主动预测与搜索技术能够实现对边境地区的实时监控和智能预警，有效提高边防安全和防御能力，可以应用于边境管控、缉私缉毒、人员管控等多个场景；其次，通过技术共享，主动预测与搜索技术还能促进我国边境全域范围内边境安全的共同提升。

(2) 创新

* 针对超大场景下变化物体特征有效提取问题，提出基于图结构的特征抽象方法，提升了复杂地形地貌下的特征提取完整性。

* 针对关键动态图结构与网络模块获取与解释问题，提出动态图迭代查询与解释方法，提升稀疏观测下的特征有效利用率和图节点重要性可解释性。

* 针对多条件优化的强化学习算法收敛慢的问题，提出多条件强化学习快速训练方法，提升强化学习代理网络的训练效率和推理质量。

5. 年度研究计划及预期研究结果（包括拟组织的重要学术交流活动、国际合作与交流计划等）。

5.1 年度研究计划

本课题将围绕超大场景安防监控背景下，变化目标主动预测与搜索过程中的几个关键技术开展研究，开展过程中，将广泛参与国内外高水平学术会议，研究成果拟投稿 ICML、ICLR、NeurIPS 等国际顶级会议和 IEEE TPAMI、Pattern

Recognition、计算机学报、软件学报、自动化学报等国内外权威期刊。研究计划在 4 年内完成，具体如下：

2026 年 1 月 1 日-2026 年 12 月 31 日

- (1) 专家咨询与资料收集：咨询边防监控领域的专家，了解监控场景的特点和需求。收集相关资料，包括超大场景监控的图像数据、监控设备的性能参数等；
- (2) 场景模拟与数据准备：拍摄实验室周围场景的视频，模拟超大场景边防监控环境。对拍摄的视频进行处理，提取图像序列，构建数据集；
- (3) 图像分割与特征图构建：研究现有的图像分割方法，如基于 CNN，BERT 和 Mamba 的分割算法。利用这些方法对模拟场景和真实边防图像序列进行分割，构建变化物体的特征图结构；
- (4) GNN 特征提取与初步分析：将构建的特征图结构输入图神经网络(GNN)，进行特征提取。分析提取的特征，为后续的目标位置预测提供基础；
- (5) 年度小结，撰写不少于 2 篇高质量论文并投稿，申请国家发明专利 1 件以上。

2027 年 1 月 1 日-2027 年 12 月 31 日

- (1) 动态图构建与分析：基于第一年构建的特征图结构序列，构建动态图。分析动态图的结构和特性，为后续的条件概率分布学习提供基础；
- (2) 条件概率分布学习：使用 GNN 学习动态图上的条件概率分布。通过训练模型，估计目标出现的可能位置；
- (3) 可信度评估：研究 GNN 的可解释性方法，如基于注意力机制的解释方法。利用这些方法评估图结构和网络模块对预测结果的可信度；
- (4) 结果分析与优化：分析预测结果和可信度评估结果，优化模型和算法；
- (5) 年度小结，撰写不少于 2 篇高质量论文并投稿，申请国家发明专利 1 件以上，参加国际或国内会议 1 次。

2028 年 1 月 1 日-2028 年 12 月 31 日

- (1) 强化学习算法设计：基于场景中目标出现位置的可信预测，设计强化学习算法，然后定义奖励函数和状态转移函数，为摄像头的控制提供决策依据；
- (2) 摄像头控制策略：使用强化学习算法训练摄像头控制策略，并通过与环境的交互，学习最优的摄像头控制策略，实现高效稳定的目标搜索；
- (3) 实验验证与优化：在模拟和真实的边防监控环境中进行实验验证，并根据实验结果，优化强化学习算法和摄像头控制策略；
- (4) 结果分析与总结：分析实验结果，总结基于强化学习的摄像头控制方法的优点和不足。
- (5) 年度小结，撰写不少于 2 篇高质量论文并投稿，申请国家发明专利 1 件以上，参加国际或国内会议 1 次。

2029 年 1 月 1 日-2029 年 12 月 31 日

- (1) 软硬件平台设计：根据前三年的研究成果，设计软硬件平台，包括数据采集模块、特征提取模块、目标预测模块、摄像头控制模块等。
- (2) 系统集成与测试：将各个模块集成到一个完整的系统中，并在实验室环境中进行初步测试，检查系统的功能和性能。
- (3) 实地测试与调优：将系统部署到边境进行实地测试。根据实际的边防监控场景，对系统进行调优，提高系统的稳定性和准确性。
- (4) 结果总结与推广：总结实地测试的结果，撰写项目总结报告，同时将研究成果进行推广，为边防监控提供新的技术手段。

5.2 预期研究结果

本项目预期取得如下研究成果：

- * 建立超大场景下动态目标主动预测与搜索的关键技术体系；
- * 构建基于图结构抽象的变化场景及目标特征提取方法、基于稀疏观测的动态图可信预测方法以及多约束条件下强化学习模型快速训练模式；

- * 开发面向边境区域超大场景主动目标预测与搜索软硬件平台 1 套；
- * 发表 CCF 推荐 A、B 类中英文期刊及会议论文 4-6 篇；
- * 申请国家发明专利 2-3 项；
- * 撰写课题研究报告 1 份；
- * 培养研究生 4-5 名。

（二）研究基础与工作条件

1. 研究基础（与本项目相关的研究工作积累和已取得的研究工作成绩）；

2. 工作条件（包括已具备的实验条件，尚缺少的实验条件和拟解决的途径，包括利用国家实验室、全国重点实验室和部门重点实验室等研究基地的计划与落实情况）；

本课题依托云南师范大学信息学院开展工作，拥有实验室用房 500 多平米，高性能图形工作站 30 余台、云计算平台 5 套、行业级无人机 5 台、大型双光监控设备 1 套，移动便携监控设备 2 套，具备了开展项目所需的工作条件。申请人目前承担云南省科技厅基础研究计划青年项目 1 项，云南省智能系统与计算重点实验室开放课题 1 项目，前期的探索与研究已经为本课题的顺利开展打下了坚实的基础。申请人所在云南师范大学模式识别与人工智能实验室主要研究计算机视觉，模式识别，图数据处理与表示学习，强化学习等机器学习领域的理论及其应用，可为本项目的开展提供多方面的支持。同时，依托云南师范大学信息学院，并与云南省内边防管理部门建立了良好的合作关系，能够为数据采集、定期监测等提供帮助。

3. 正在承担的与本项目相关的科研项目情况（申请人和主要参与者正在承担的与本项目相关的科研项目情况，包括国家自然科学基金的项目和国家其他科技计划项目，要注明项目的资助机构、项目类别、批准号、项目名称、获资助金额、起止年月、与本项目的关系及负责的内容等）；

无

4. 完成国家自然科学基金项目情况（对申请人负责的前一个资助期满的科学基金项目（项目名称及批准号）完成情况、后续研究进

展及与本申请项目的关系加以详细说明。另附该项目的研究工作总结摘要（限 500 字）和相关成果详细目录）。

无

（三）其他需要说明的情况

1. 申请人同年申请不同类型的国家自然科学基金项目情况（列明同年申请的其他项目的项目类型、项目名称信息，并说明与本项目之间的区别与联系；已收到自然科学基金委不予受理或不予资助决定的，无需列出）。

无

2. 具有高级专业技术职务（职称）的申请人或者主要参与者是否存在同年申请或者参与申请国家自然科学基金项目的单位不一致的情况；如存在上述情况，列明所涉及人员的姓名，申请或参与申请的其他项目的项目类型、项目名称、单位名称、上述人员在该项目中是申请人还是参与者，并说明单位不一致原因。

无

3. 具有高级专业技术职务（职称）的申请人或者主要参与者是否存在与正在承担的国家自然科学基金项目的单位不一致的情况；如存在上述情况，列明所涉及人员的姓名，正在承担项目的批准号、项目类型、项目名称、单位名称、起止年月，并说明单位不一致原因。

无

4. 同年以不同专业技术职务（职称）申请或参与申请科学基金项目情况（应详细说明原因）。

无

5. 其他。

无