

# 基于空间特征选择的自适应跟踪算法

江南大学 徐天阳 吴小俊 萨里大学 冯振华 Josef Kittler

复杂环境下视频目标跟踪旨在自动建立目标及其周围背景的动态表观模型，实现自适应地目标定位，对于视觉环境理解、行为分析、视频检索等问题具有重要意义。近年来，通过深度神经网络得到的物体表观描述信息具有更好的鉴别性。基于相关滤波算法的跟踪器通过循环矩阵实现样本扩展和频率域内的高效求解，在目标跟踪领域内得到广泛研究并取得领先效果。但现有学习算法仅考虑不同空间位置固定权重，并没有将自适应空间特征选择考虑在内，往往容易受到干扰信息影响，对最终目标位置的推断产生影响。

基于以上考虑，提出了一种基于空间特征选择的自适应跟踪算法。该算法强调空间特征选择对目标跟踪的重要性，尤其是背景中部分高对比空间信息可以帮助实现鲁棒的目标跟踪。将自适应空间特征选择与相关滤波算法框架结合的优点如下。1) 通过对高维特征进行自适应空间特征选择，可实现对滤波器的降维学习，使其在跟踪过程中处于低维流形上；2) 通过对高维视觉特征进行空间区隔，可以实现目标和周围背景的互补学习，支撑当前鉴别性空间位置将被激活，形成稳定空间支撑模式，充分利用了前景背景中互补信息；3) 通过对不相关空间位置抑制，大大减少干扰信息，增强了算法鲁棒性。

如图 1 所示，在目标和背景都存在复杂变化

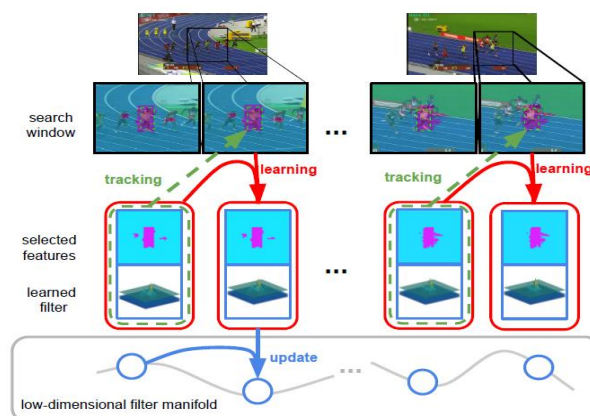


图 1 基于空间特征选择的自适应跟踪算法

的情况下，基于空间特征选择的自适应跟踪算法能够学习到合适的空间支撑模式，将前景和背景信息中的互补信息进行协同考虑。通过将原本的背景干扰物和目标进行协同考虑，增强了对目标本身的支持，同时也能够在目标变化过大的情况下依靠背景信息对目标位置进行合理推断。

上述算法在诸多目标跟踪数据集如 VOT2018, OTB100, TC128 和 UAV123 都取得了较高的跟踪精度和鲁棒性。该研究的相关成果在 VOT2018 竞赛中取得了公开数据库第一名的成绩（代码下载：<https://github.com/XU-TIANYANG/LADCF>），并被国家自然科学基金委报导“视觉跟踪重要进展”（<http://www.nsf.gov.cn/publish/portal0/tab434/info74717.htm>）。（责任编辑：王金甲）



徐天阳

江南大学在读博士研究生，目前在英国萨里大学计算机视觉，语音和信号处理中心进行为期一年的联合培养。主要研究领域包括视觉目标跟踪和机器学习。

Email: [tianyang.xu@surrey.ac.uk](mailto:tianyang.xu@surrey.ac.uk)



吴小俊

江南大学二级教授，博导，江南大学科学技术研究院院长，江苏省人工智能学会副理事长。主要研究方向为模式识别、计算机视觉、智能计算和机器学习。

Email: [wu\\_xiaojun@jiangnan.edu](mailto:wu_xiaojun@jiangnan.edu)

# 融合单阶段和两阶段检测思想的物体检测算法

中科院自动化所 张士峰 雷震 李子青

物体检测是视觉感知的第一步，也是计算机视觉的一个重要分支。物体检测的目标是用框去标出物体的位置，并给出物体的类别。如图 1 所示，目前基于深度学习的物体检测算法主要分为单阶段方法和两阶段方法。通常情况下，单阶段方法检测速度更快，而两阶段方法检测精度更高。

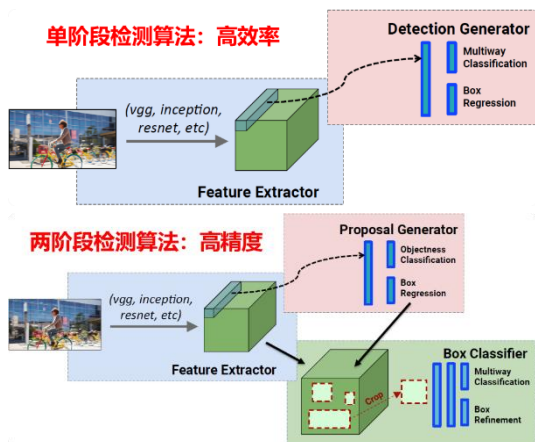


图 1 单阶段和两阶段检测算法的对比

为了继承两种方法各自优点并克服它们缺点，我们提出了 RefineDet 通用物体检测算法。该算法融合了单阶段和两阶段方法的检测思想，在准确率和速度上达到了很好的均衡。RefineDet 算法是在网络结构上对单阶段方法进行改进，以模拟两阶段方法的检测过程，从而具备两阶段方法优势并能够保持单阶段方法速度。

如图 2 所示，RefineDet 算法由锚点框校正模块和物体检测模块组成，它们之间由传输连接

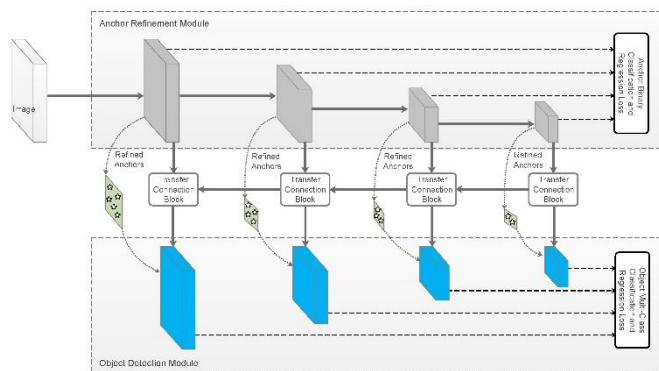


图 2 RefineDet 算法的网络架构示意图

块连接。锚点框校正模块旨在滤除简单的负样本以减少后续分类器的搜索空间，同时粗略地调整锚点框的位置和大小以为后续回归器提供更好的初始化。物体检测模块把经过前者矫正过的锚点框作为输入，以进一步提高回归精度并预测多类别标签。由于锚点框校正模块中的特征被用于二分类，已具备一定鉴别能力，传输连接块把锚点框校正模块中这些有鉴别力的特征，传递给物体检测模块，以更好地预测物体位置，大小和类别标签。通过这种网络结构，RefineDet 能够在保持单阶段方法的速度前提下，具备两阶段方法的二阶段分类、二阶段回归、二阶段特征这三个优点。因此该算法能够在常用的通用物体检测数据库上取得最好检测精度的同时，保持着较快检测速度，在准确率和速度上达到了很好均衡。该研究成果发表在计算机视觉国际会议 CVPR 2018。

(责任编辑：崔海楠)



张士峰

中科院自动化所博士。主要研究方向为物体检测。

Email: shifeng.zhang@nlpr.ia.ac.cn



雷震

中科院自动化所研究员。主要研究方向为生物特征识别，计算机视觉。

Email: zlei@nlpr.ia.ac.cn



李子青

中科院自动化所研究员，IEEE Fellow。主要研究方向为生物特征识别，计算机视觉。

Email: szli@nlpr.ia.ac.cn

# 基于双链区分性卷积神经网络的多视角 3D 模型检索算法

山东省科学院 山东省计算中心（国家超级计算中心） 高赞

随着传感器技术、存储技术和网络传输技术的快速发展，人们获得 3D 模型的途径也越来越多。随着 3D 模型应用领域的增广，通用模型数据库中 3D 模型的种类及数目也随之增长，如何快速精确查找所需的 3D 模型也逐渐成为一个热门研究话题。然而，3D 模型不仅受到光照、尺度、物体形变和对象之间差异等因素影响，而且其本身复杂性也给 3D 模型特征描述带来了很大困难，这些困难给 3D 模型检索带来了巨大挑战。

近年来，研究人员提出了很多不同的 3D 模型检索算法，其中基于多视角的 3D 模型算法的性能表现尤为突出（将 3D 模型虚拟投影为多视角图像），该类算法大致可以分为两类：基于手工设计特征的算法和基于特征学习的算法。基于手工设计特征的算法首先通过手工设计特征对每张图像进行特征提取，然后，通过不同的检索算法计算图像之间的相似程度。由于特征提取和检索过程不统一，特征的区分性和检索性能都有待提高。基于特征学习的算法将特征提取和检索过程统一到同一个框架中，其所学习的特征更适合于后续的检索过程。这类方法的成功关键在于：1) 如何自动地挖掘多视角样本潜在关联特性？2) 如何提高所提取特征的区分性？3) 如何高效地对模型进行训练？

针对这一问题，我们提出了基于双链区分性卷积神经网络的多视角 3D 模型检索算法。该方法通过结合对比损失和对比中心损失函数，从训

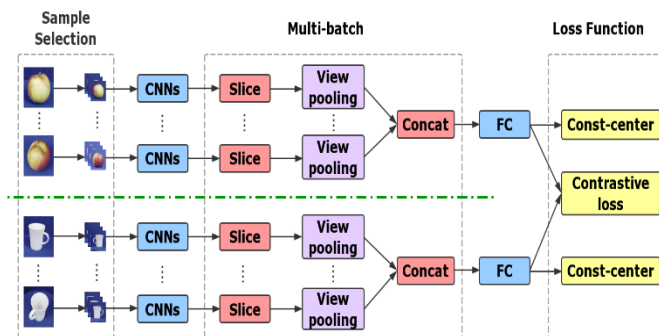


图 1 双链区分性卷积神经网络的框图

练样本中，自动选择性地蒸馏出最富信息的训练样本，高效地对网络进行训练，提高所提取特征的区分性，同时，通过加入切分层和连接层，使得网络能够同时输入多视图和多个 Batch 的训练样本，从而挖掘了多视角样本的关联特性，加快网络的收敛速度和提高模型的检索性能。如图 1 所示，该算法一共包含两条链，分别对应的是 Query 链和 Gallery 链。在训练过程中，两条链的输入样本均来自于训练样本，且两条链的输入均为经过蒸馏的多视角样本对。通过该端到端的双链网络，不仅能够减少对原始数据集中带标签训练样本的要求，而且能够自动挖掘多视角样本的潜在关联特性，同时，使得所学习特征具有更好的类内和类间的特性，从而增强了所提取特征的区分性，提高了 3D 模型检索的性能。

该研究相关成果发表在 IEEE CVPR、AAAI 和 ACM MM 上。

（责任编辑：任桐炜）



高赞

山东省科学院 山东省计算中心（国家超级计算济南中心），研究员，博士生导师，主要研究方向为人工智能、图像/视频分析、信息检索和机器学习及其应用。Email: gaoz@sdas.org