
软件所在智能理论研究中取得进展

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/10857.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

近日，中国科学院软件研究所智能软件研究中心在人工智能基础理论方面取得进展，在树形神经网络、风格迁移模型、小目标计数具体问题上提出新的解决方案。研究成果发表在人工智能领域会议CVPR 2020、ECCV 2020、ACM MM2020，并受邀完成线上学术报告，相关代码已经开源。相关研究包括：

注意力卷积二叉神经树

细粒度分类的目的是对粗粒度的大类别进行更加细致的子类划分。通常情况下，细粒度数据集中预先给定的标注往往依赖于专业人工标注，需专业的领域知识，所以标注的数量很少并且很难获得。关键的判别力区域散布在细小区域，需要网络进行定位的工作，而现有大多数的网络将定位的工作与识别整合在一起，网络缺乏捕获判别力区域的能力。因此，亟待寻求能够学习有效特征和区分难易特征的方法。

针对这一问题，软件所团队将卷积神经网络与软决策树策略相结合，提出一种二分类深度神经树的细粒度图像识别模型，该模型包括主干模型（backbone）、路由（router）、转换（transformer）、预测（predictor）四种模块构成。在特征提取的主干网络上，选取VGG-16作为团队的特征提取器，这一部分包括VGG-16的conv1-conv5的部分卷积层。在浅层利用路由进行二分类的预测输出，将图像预先做粗粒度分类，随着树的深度增加，类别的区分范围不断缩小，并在叶子节点处进行最终类别的输出。借助这种由粗分逐渐到细分的方法，团队在神经树的边上构造了由卷积算子构成的变换模块，在不同的分支上利用卷积网络进行特征学习，定位出具有判别力的区域特征，使每个叶子节点上的预测器重点关注图像不同的区域。为加强不同分支对图像多尺度特征的提取，团队采取非对称策略来安排转换模块在不同分支上的数量，而不同的分支之间起到了相互补充与促进特征学习的作用。在预测阶段，团队利用路由模块预测值的累积作为该路径被选择的概率，相应的叶子节点的输出则是该类别的预测概率分布。

实验结果显示，在CUB-200-2011、Stanford Cars和Aircraft数据集上，基于深度神经树的细粒度识别模型显著优于当前其他的弱监督细粒度方法。相关成果以Attention Convolutional Binary Neural Tree for Fine-Grained Visual Categorization为题，被CVPR 2020收录。（[代码链接](#)）

语义神经树

现有的人体解析方法往往忽略人体结构中各部位之间固有的语义关系，把这种细粒度的分割任务看作常规的语义分割，这样的操作易引起细小部位分割错误。此外，人体部位之间的尺度也存在相当大的差异，如人的双脚、袜子等一般在图像中占有很小比例甚至在图像中不出现。

针对这一问题，软件所团队提出一种基于深度神经树来学习语义特征。多叉深度神经解析树采用了一种层次关系，在浅层中，团队先把图像中的前景与背景进行识别，利用产生的掩码信息作用在输入数据上，产生新的前景信息将输送到下一层模块处理，在不同的层级上，团队会将相应的部位识别并分割出来。此外，团队还构造了融合模块用来收集在不同的层级产生的结果，并将结果进一步优化调整，产生最终的分割预测结果。

多叉神经语义解析树利用随机梯度下降（SGD）方法以端到端的方式进行训练整个模型，具有收敛速度快，性能稳定的特点。在LIP、PASCAL-Person-Part、CIHP和MHPv2四个数据集上进行的单人和多人解析实验，均证明了该方法的有效性。相关成果以Learning Semantic Neural Tree for Human Parsing为题，被ECCV 2020收录。（[代码链接](#)）

无监督域适应的空间注意金字塔网络

域适应是解决目标域和源域数据分部有显著差异的问题。巧妙进行迁移转化，能够显著改善在目标域的实际使用效果。

针对这一问题，软件所团队提出空间注意力金字塔机制，将域转换的注意力集中在与任务相关的特定区域。此方法将特征图用不同大小的窗口来提取均值特征，获得特征金字塔，然后用软注意力机制在不同金字塔特征之间的动态选择，来指导最终的域转换。

实验结果显示，提出的方法在多个任务（目标检测、实例分割、语义分割）中均取得领先的效果。相关成果以Spatial Attention Pyramid Network for Unsupervised Domain Adaptation为题，被ECCV 2020收录。（[代码链接](#)）

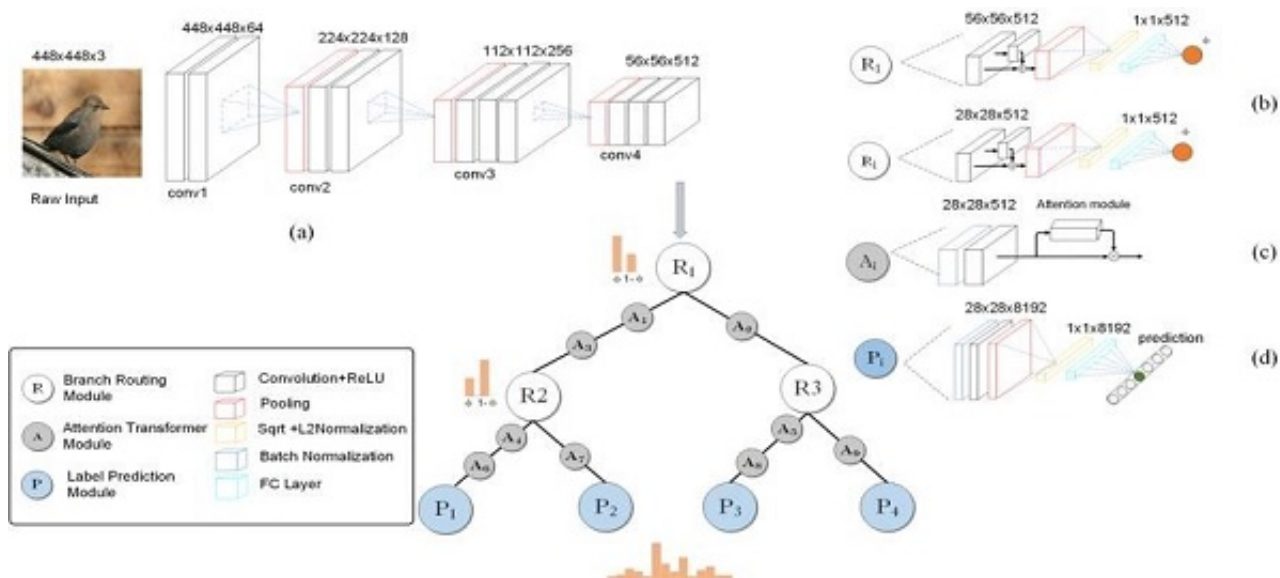
用于无人机目标检测和计数的引导注意力网络

检测和计数是人工智能的基础问题。在无人机场景中，受目标小、背景杂乱等复杂因素影响，通过无人机采集视频进行目标检测和计数具有极大的挑战。

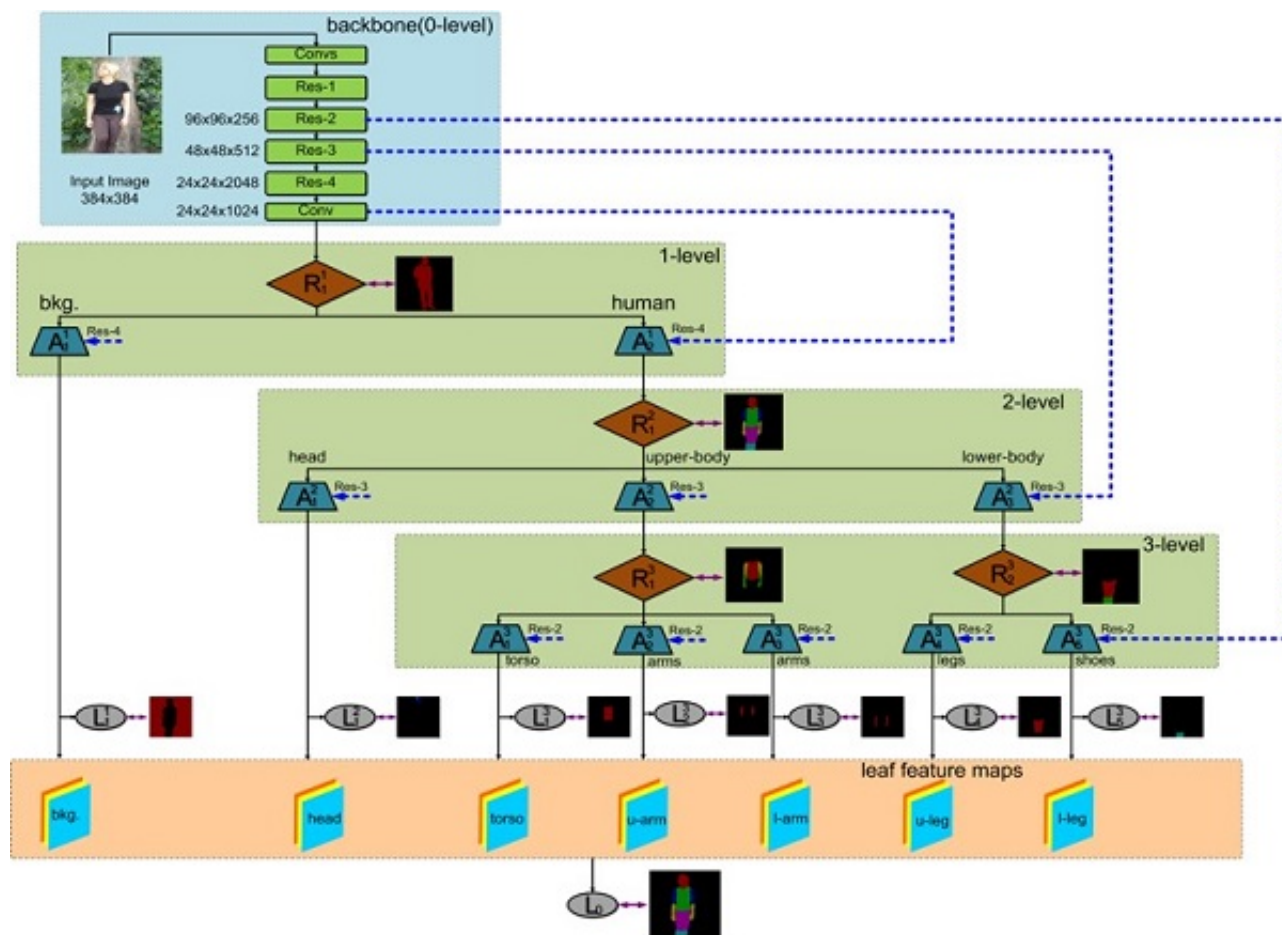
针对复杂场景的小目标检测和计数问题，软件所团队设计了一种新的基于特征金字塔的注意力网络来处理目标检测和计数任务。与以往依赖于无监督注意模块的方法不同，团队充分利用了背景和目标之间的弱监督信息，融合不同尺度的特征图，获得更语义化的特征表示。同时，前景模块考虑了全局和局部外观的目标，以方便准确定位。

团队设计了新的数据增强策略，用于在各种复杂场景下训练鲁棒模型，在UAVDT、CARPK和PU CPR+三个数据集上显示了算法先进的检测和计数性能。相关成果以Guided Attention Network for Object Detection and Counting on Drones为题，被ACM MM 2020收录。（[代码链接](#)）

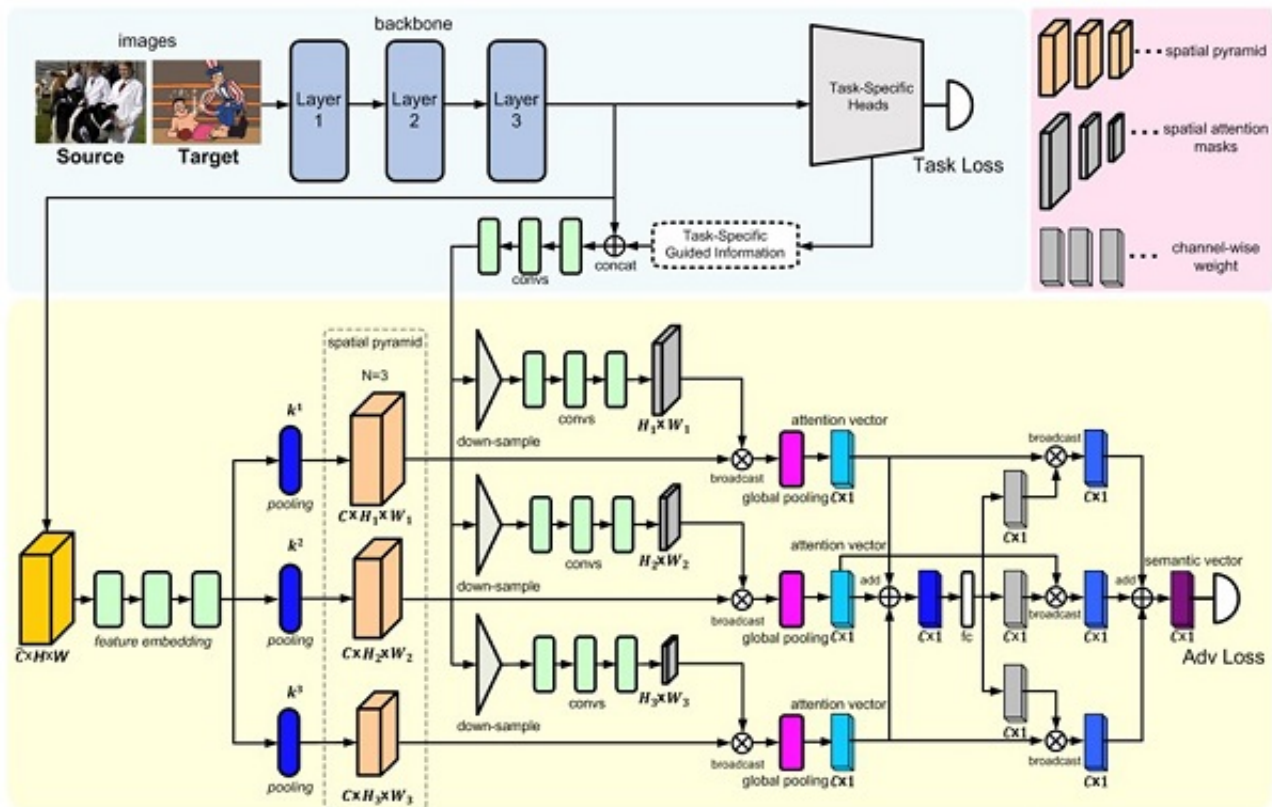
以上成果论文通讯作者为软件所副研究员张立波。相关研究得到中科院青年创新促进会、软件所优秀青年科技人才计划支持。



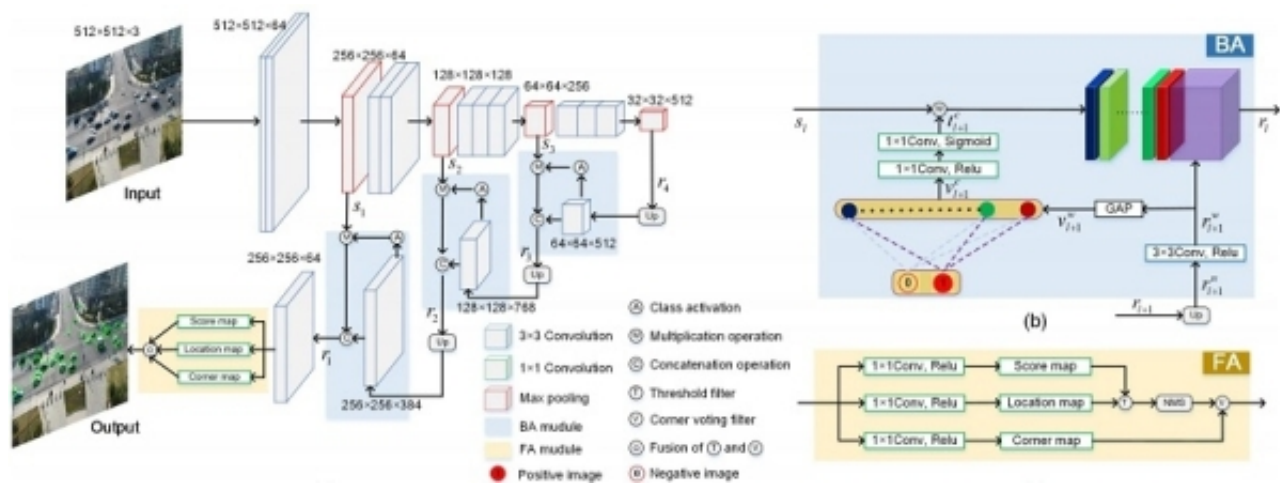
卷积二叉神经树模型



基于语义神经树的人体解析模型



空间注意金字塔网络



基于引导注意力机制的无人机检测和计数模型

研究团队单位：软件研究所

更多科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](http://iikx.com)转发