



探索发现 · 交大智慧

上海交大卢策吾团队实时多人姿态估计系统升级，挑战拥挤人群场景

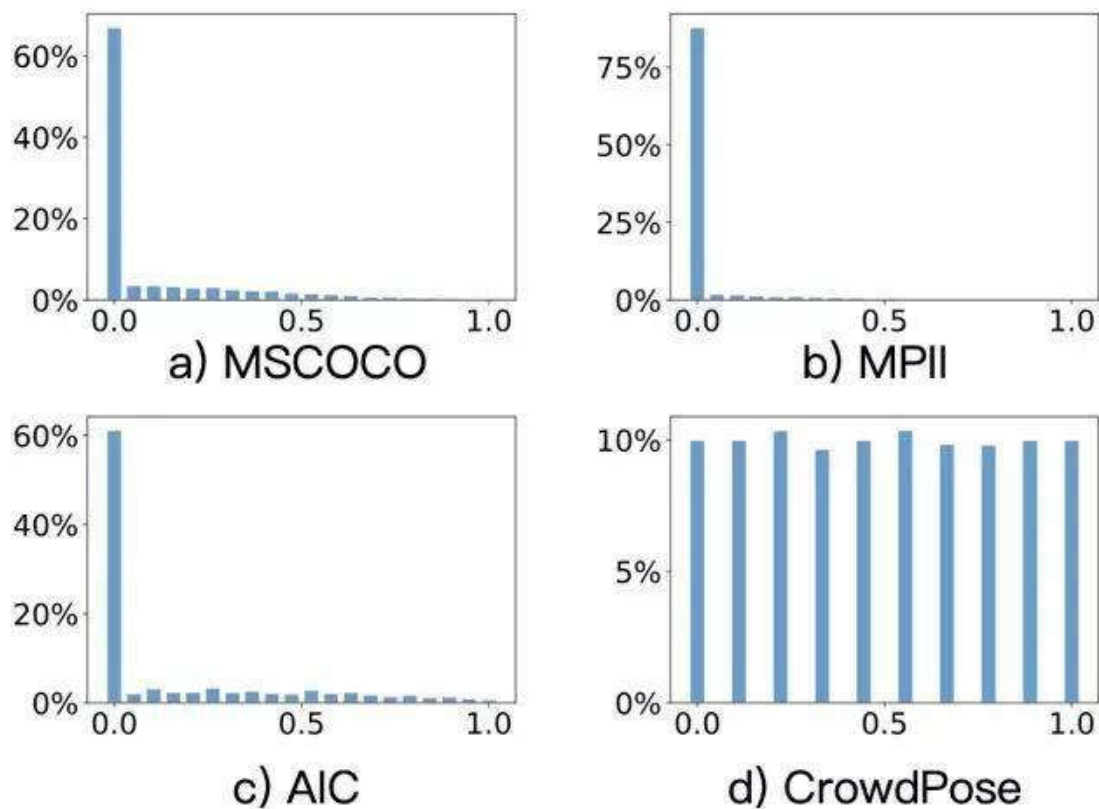
2019年03月08日 责任编辑：洪纳娜 孙佳



近日，上海交通大学电子信息与电气工程学院卢策吾团队（MVIS组）的论文《CrowdPose: Efficient Crowded Scenes Pose Estimation and A New Benchmark》被CVPR2019（计算机视觉领域顶级会议）接收。该论文主要是针对此前AlphaPose系统的优化，构建了CrowdPose数据集，用来衡量算法在拥挤场景中的性能，并提出了一个高效的算法来解决拥挤人群中的姿态估计问题，实验结果比起当前最好的算法有较大提高。

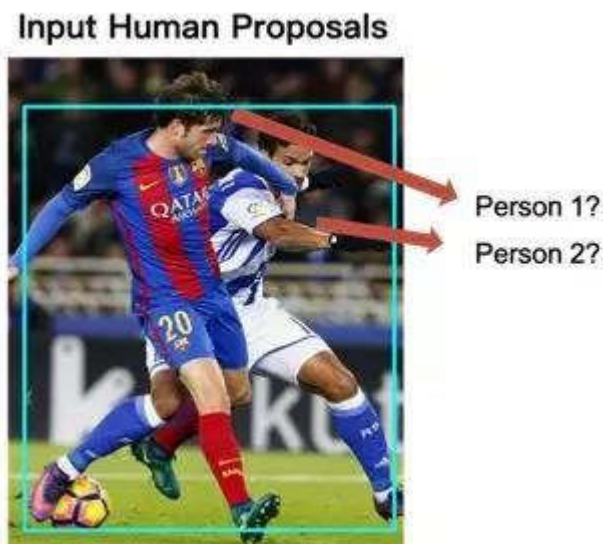
为了衡量人体姿态估计算法的性能，学术界与工业界建立了越来越多的公开数据集，如MPII、MSCOCO、AI Challenger。然而，这些数据集的图片通常采集于日常生活场景，缺乏拥挤人群场景的数据。对于由数据驱动的深度学习方法来说，数据集的分布不均衡，意味着算法性能的不均衡。同时，随着学术界对人体姿态研究的深入，算法追求像素级的精度，简单场景下的性能已经逐步逼近人类的精度，但在拥挤场景中往往会失效。

针对以上的情况，上海交大MVIG组的研究者开源了CrowdPose数据集。CrowdPose数据集中的图片，有着均匀分布的拥挤程度。既可以评估算法在日常非拥挤情况下的表现，也可以评估其在极度拥挤时的性能。目前，数据集开源了2万张图片，评估工具也已经上线。在未来的几个月时间内，研究者将会不断扩大数据的规模，开源一个更大的数据集。



三大主要数据集与CrowdPose数据集的拥挤场景数据分布对比

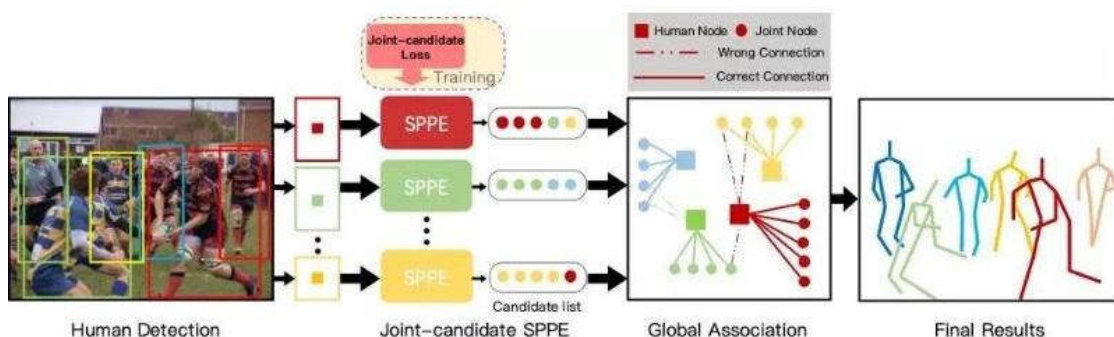
就算法而言，在拥挤人群的场景下，由于人群过于密集，重合程度太高，每个人的位置难以用人体检测框表示，传统的二步法模型往往会失效。因此，研究者们提出了一个全局竞争匹配算法，减少了姿态估计模型对于人体框的依赖，同时提高了模型对于复杂人体场景的鲁棒性，在拥挤场景中的表现超越了现有的方法。



人体检测框难以表达人的位置

研究者们设计了一个关节候选损失函数，通过控制模型输出响应程度不同的多峰值热度图，在人体框不准确的情况下，模型尽可能地输出候选关节可能的位置。在得到每个人体框的候选关节后，通过聚类与链接，消除冗余结果的同时，构建出一个人体实例-候选关节的图模型，以此来表征了人体实例与每个关节点之间的连接关系与概率。

借助人体实例-候选关节的图模型，人体姿态估计问题可转化成图模型中的最佳匹配问题，因为人体姿态问题具有很强的特殊性，而该图模型具有很强的稀疏性。并且经过研究者的分析，优化匹配问题的时间复杂度与传统的NMS算法相当。



本研究全局竞争匹配算法概览

总得来说，该方法由于建立了一个全局的图模型，在匹配过程中考虑到了整体的连接方式，因此能很好地改善了二步法中缺乏全局视野的不足。

此外，在实验中，研究者们还对比了该算法与其他开源系统与算法之间的性能。在CrowdPose数据集中，比当前最好的算法提升了5.2mAP；在极度拥挤的子集中，提升了6.2mAP。相较于OpenPose和Detectron（Mask R-CNN）等姿态估计开源系统，不仅性能得以提升，而且运行速度也快了2至3倍。

Method	mAP @0.5:0.95	mAP@0.5	mAP @0.75	mAR @0.5:0.95	mAR @0.5	mAR @0.75
Mask R-CNN [14]	57.2	83.5	60.3	65.9	89.5	69.4
AlphaPose [13]	61.0	81.3	66.0	67.6	86.7	71.8
Xiao <i>et al.</i> [32]	60.8	81.4	65.7	67.3	86.3	71.8
AlphaPose/Crowd	66.0	84.2	71.5	72.7	89.5	77.5

Table 1. Results on CrowdPose test set.

在CrowdPose数据集上的定量对比实验

Method	AP_{Easy}	AP_{Medium}	AP_{Hard}	FPS
OpenPose [5]	62.7	48.7	32.3	5.3
Mask R-CNN [14]	69.4	57.9	45.8	2.9
AlphaPose [13]	71.2	61.4	51.1	10.9
Xiao <i>et al.</i> [32]	71.4	61.2	51.2	-
AlphaPose/Crowd	75.5	66.3	57.4	10.1

Table 2. Results on CrowdPose test set. Test set is divided into three part and we report the results respectively. FPS column reports the runtime speed on the whole test set.

三个场景子集中的性能评估与运行速度对比（简单、一般、拥挤）

未来，研究者们将针对这种新的竞争匹配模型进行优化，对其他形式的拥挤检测问题（如拥挤实例分割）开展进一步研究。

作者： 李杰锋

供稿单位： 电子信息与电气工程学院

沪ICP备05052060 沪举报中心 版权所有© 上海交通大学 新闻网编辑部维护

地址：上海市东川路800号 邮编：200240 查号：86-21-54740000