中图分类号：TP391;TP183　**文献标识码：**A　**文章编号：**
论文引用格式：（论文引用格式：）[DOI:]

结合密集连接的轻量级高分辨率人体姿态估计

高坤[1]，李汪根[1]，束阳[1]，葛英奎[1]，王志格[1]

1.安徽师范大学计算机与信息学院，安徽省芜湖市 241002

摘　要：目的 为了更好地实现轻量化的人体姿态估计，在轻量级模型极为有限的资源下实现更高的检测性能。**方法** 基于高分辨率网络（HRNet）提出了结合密集连接网络的轻量级高分辨率人体姿态估计网络（LDHNet）。通过重新设计 HRNet 中的阶段分支结构以及提出新的轻量级特征提取模块，构建了轻量高效的特征提取单元，与此同时对多分支之间特征融合部分进行了轻量化改进，进一步的降低模型的复杂度，最终大幅降低了模型的参数量与计算量，实现了轻量化的设计目标，并且保证了模型的性能。**结果** 通过实验表明，在 MPII 测试集上相比于最轻量的自顶向下的轻量级人体姿态估计模型 LiteHRNet，LDHNet 仅通过增加了少量参数量与计算量，平均预测准确度即提升了 1.5 个百分点，与 LiteHRNet 的改进型 DiteHRNet 相比也提升了 0.9 个百分点，在 COCO 验证集上的结果表明，与 LiteHRNet 相比，LDHNet 的平均检测准确度提升了 3.4 个百分点，与 DiteHRNet 相比也提升了 2.3 个百分点，与融合 Transformer 的 HRFormer 相比，LDHNet 在参数量更低并且计算量相当的条件下有着近似的检测性能，在面对实际场景时 LDHNet 也有着稳定的表现，在同样的环境下 LDHNet 的推理速度要高于基线 HRNet 以及 LiteHRNet 等。**结论** 以上实验结果表明该模型有效实现了轻量化并保证了预测性能。
关键词：人体姿态估计；轻量级网络；密集连接网络；高分辨率网络；多分支结构

Lightweight high-resolution human pose estimation combined with densely connected networks

Gao Kun[1]，Li Wanggen[1]，Shu Yang[1]，Ge Yingkui[1]，Wang Zhige[1]

1.School of Computer and Information, Anhui Normal University, Wuhu 241002, China

Abstract: Objective Human pose estimation is a technology that can be widely used in life. In recent years, many excellent high-precision methods have been proposed, but they are often accompanied by very large model scale, which will encounter the problem of computing power bottleneck in application. Whether for model training or deployment, large models require a lot of computing power as the basis. Most of them have low computing power. Similarly, for the scenes in daily life, the equipment needs more applicability and detection speed of the model, which is difficult to achieve by large models. Because of such requirements, lightweight human pose estimation has become a hot research field. The main problem is how to achieve higher detection accuracy and faster detection speed under the extremely limited number of resources. Lightweight models will inevitably fall into a disadvantage in detection accuracy compared with large models, but fortunately, from many studies in recent years, the lightweight model can also achieve higher detection accuracy. A good balance can be reached between them.

Method Based on High-Resolution Network (HRNet), a lightweight high-resolution human pose estimation network combined with dense connection network (LDHNet) was proposed. Firstly, by redesigning the stage branch structure in HRNet and proposing a new lightweight feature extraction module, dense connection and multi-scale were integrated to construct a lightweight and efficient feature extraction unit. By stacking multi-layer feature extraction modules and fusing the output of each layer together, the feature extraction module is composed of modules similar to the pyramid structure, and the dilated

收稿日期：；　修回日期：
基金项目：国家自然科学基金（61976006）
Supported by: The National Natural Science Foundation of China