Lite Pose: Efficient Architecture Design 2D for Human Pose Estimation

Yihan Wang[1], Muyang Li[2], Han Cai[3], Wei-Ming Chen[3], Song Han[3]
[1]Tsinghua University [2]CMU [3]MIT

CVPR JUNE 19-24 2022 NEW ORLEANS • LOUISIANA

Code: https://github.com/mit-han-lab/litepose

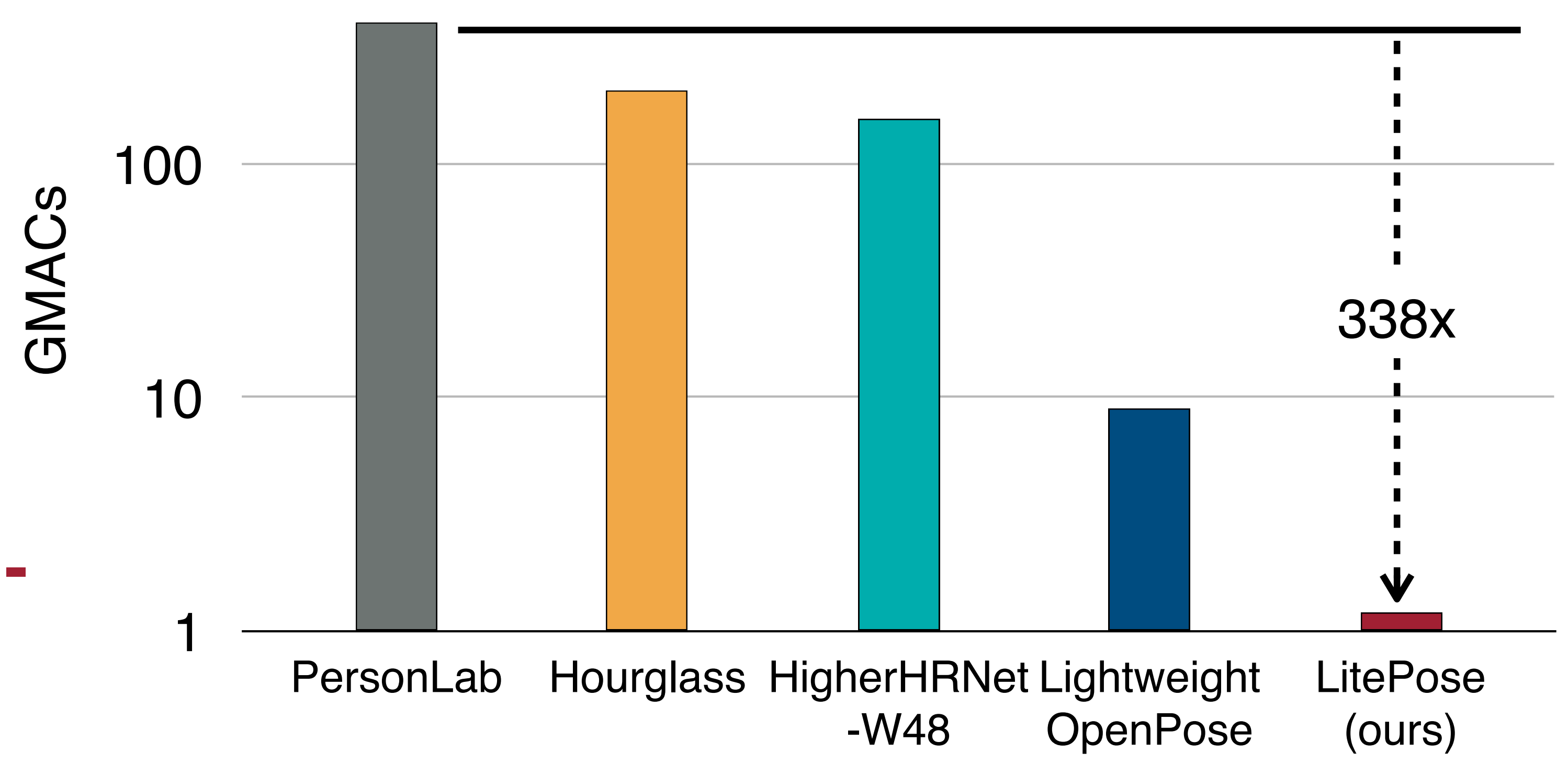
Real-Time Multi-Person Pose Estimation on Edge

Multi-Person Pose Estimation

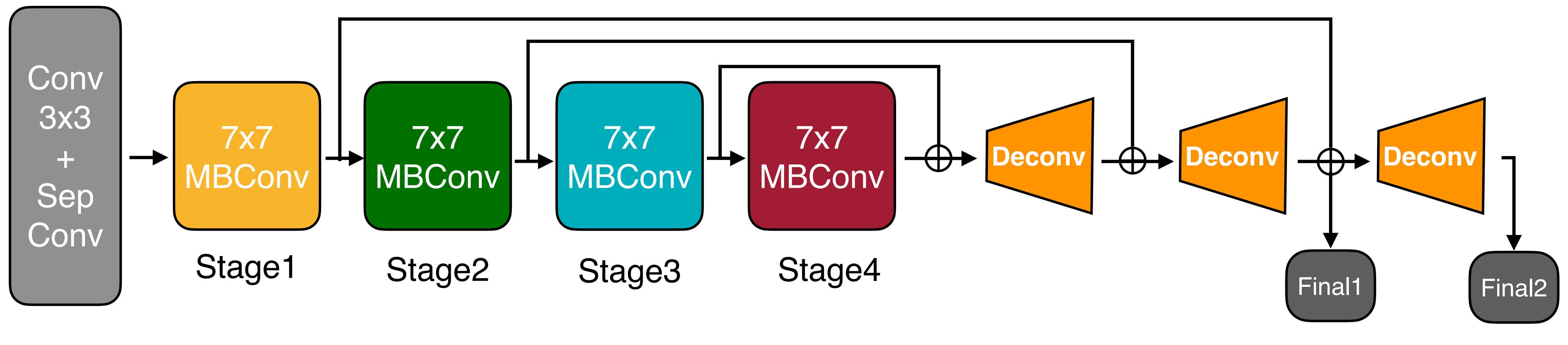
Edge Devices

Many human-centered vision applications rely on **real-time multi-person** pose estimation on **edge** devices, requiring **low-computation** pose estimation models.

However, current pose estimation models are too **heavy** for edge devices. We introduce **LitePose** to close the gap.



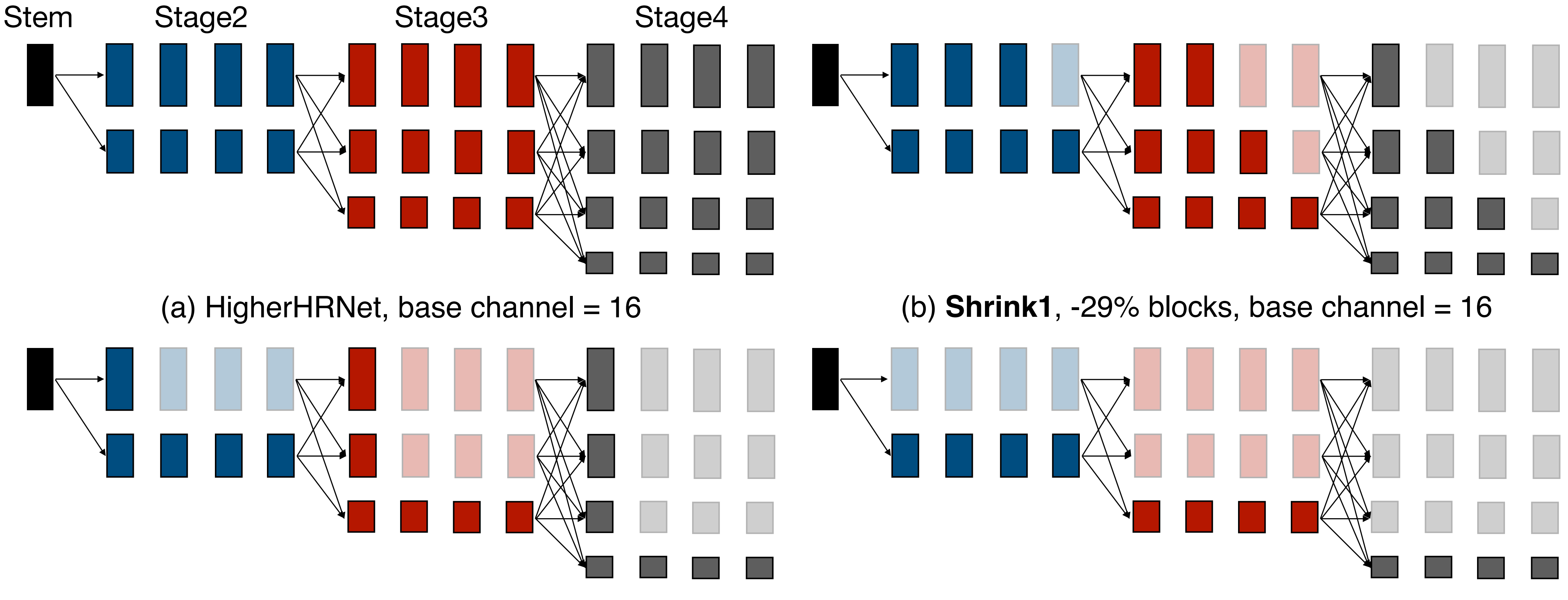
Overview of LitePose



Key insights:

1. **Single-branch architecture is efficient**
2. **Large kernel convolution is efficient.**
3. **Light-weight fusion deconv head.**

Redundancy in High-Resolution Branches

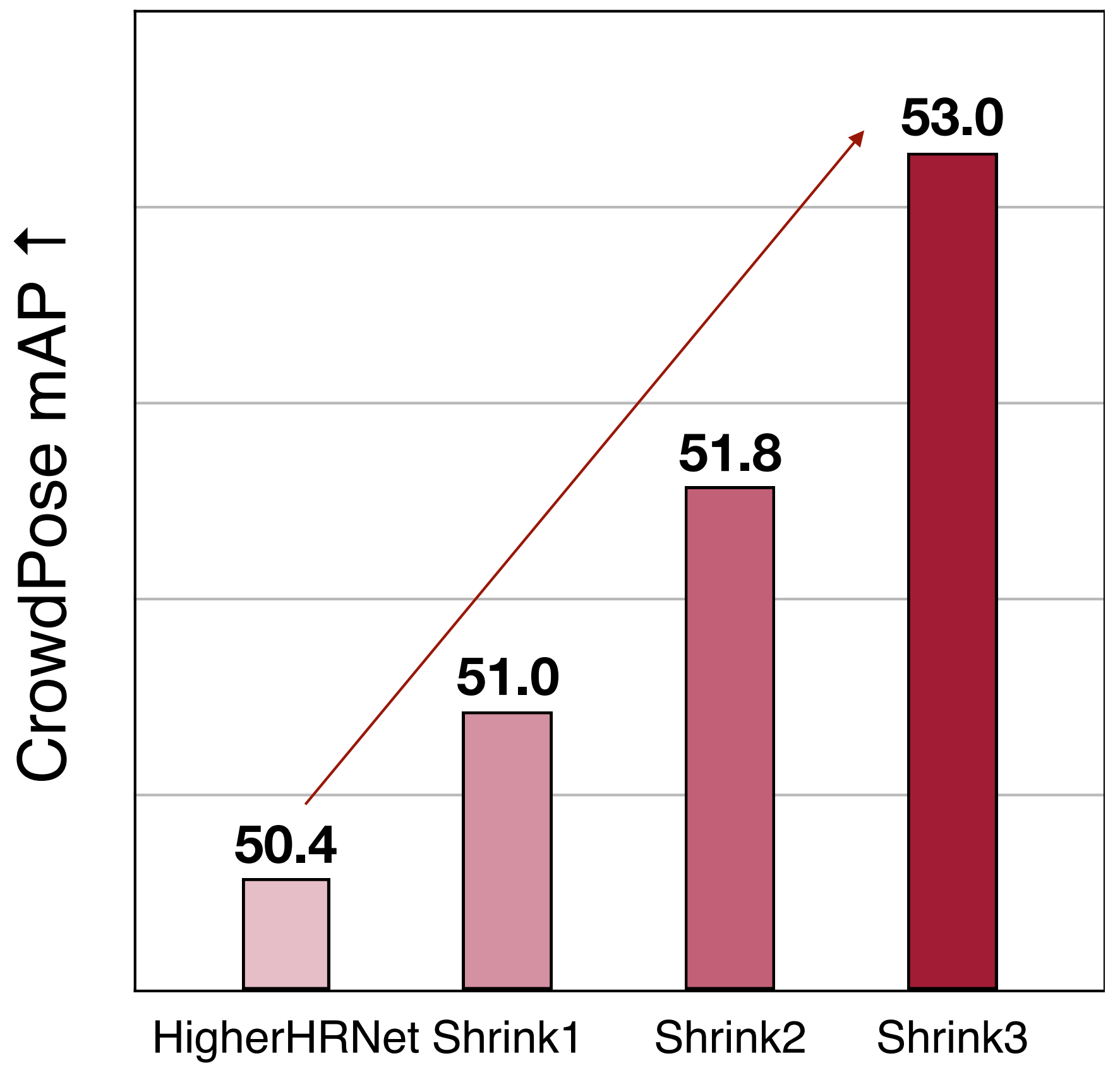


(a) HigherHRNet, base channel = 16

(b) **Shrink1**, -29% blocks, base channel = 16

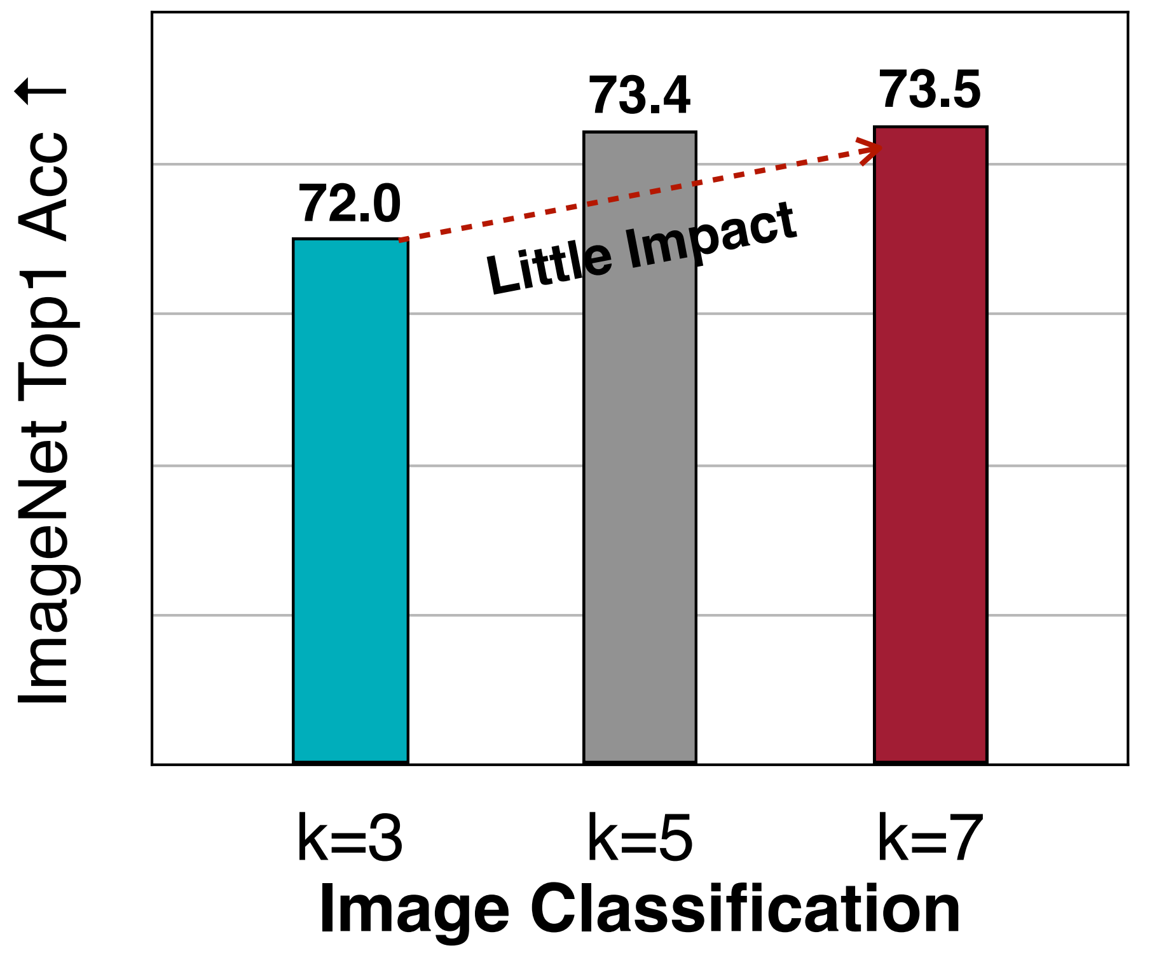
(c) **Shrink2**, -30% blocks, base channel = 18

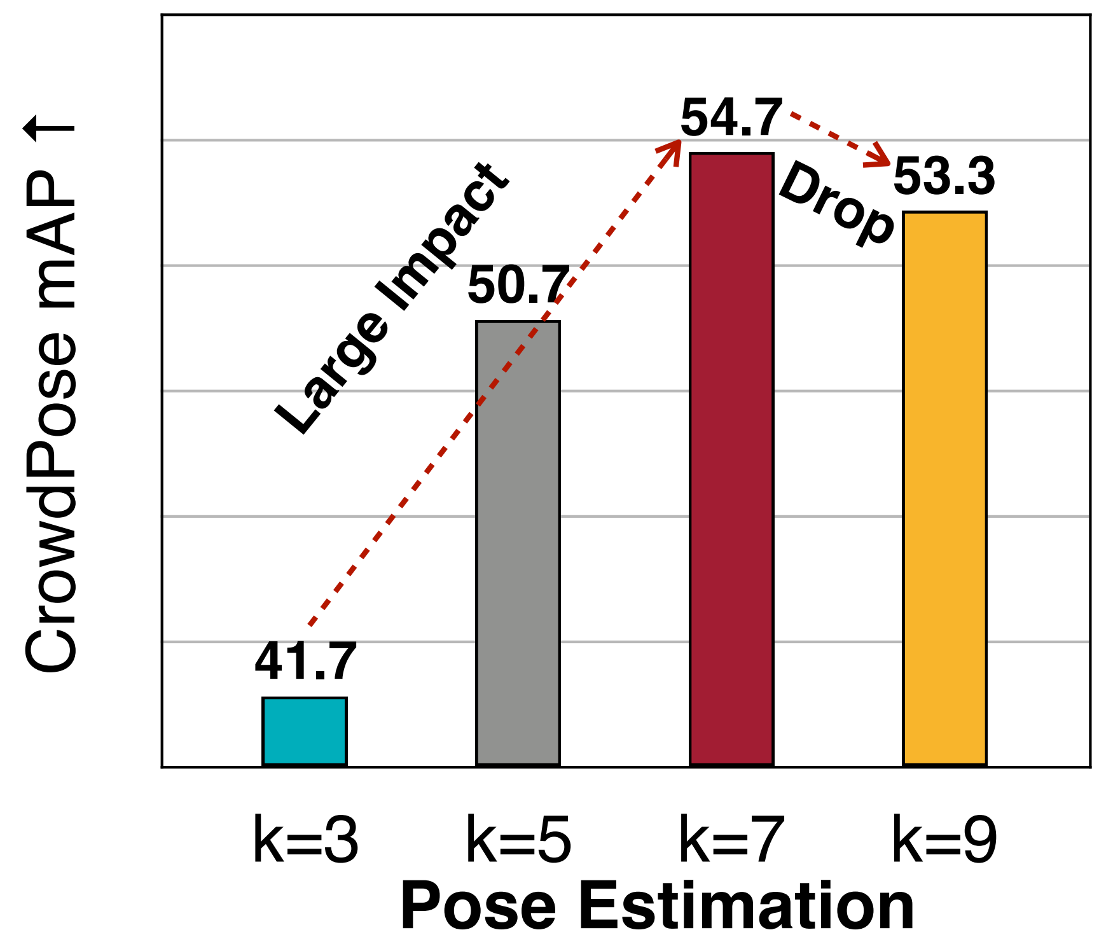
(d) **Shrink3**, -33% blocks, base channel = 18



We gradually remove blocks in high-resolution branches starting from HigherHRNet. Removed blocks are shown in transparent. The performance improves as we shrink the high-resolution branches.

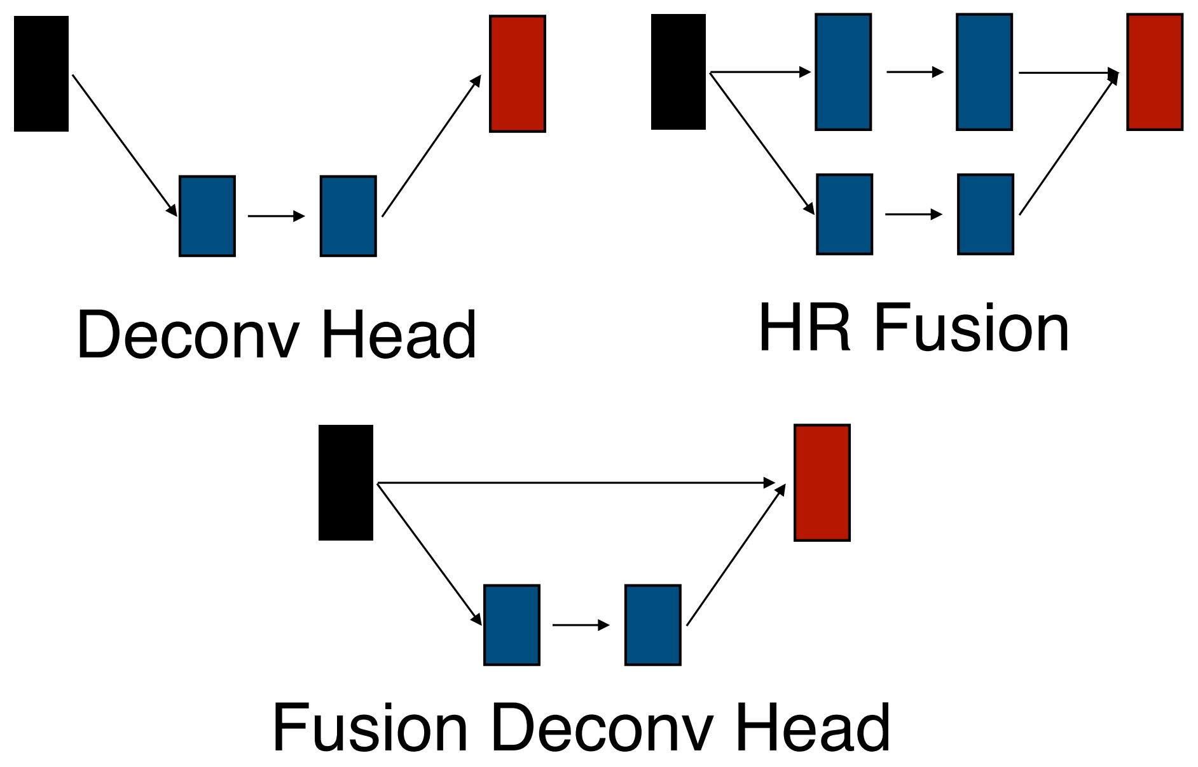
Large Kernel is Efficient

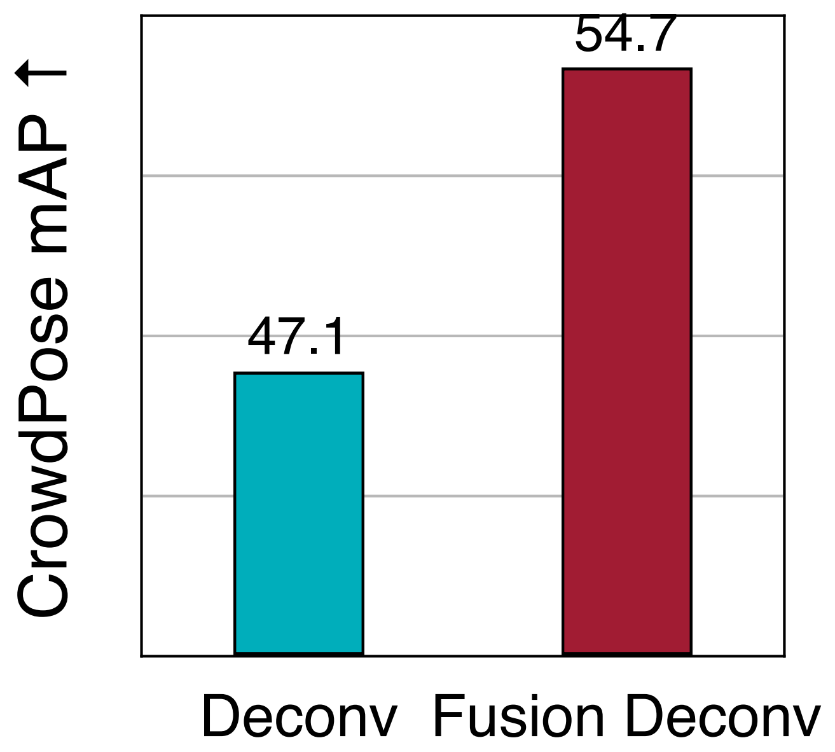




Unlike image classification, large kernel depthwise convolution plays a critical role in pose estimation. Increasing the kernel size from 3 to 7 improves the mAP by 13% mAP on the CrowdPose dataset with little overhead.

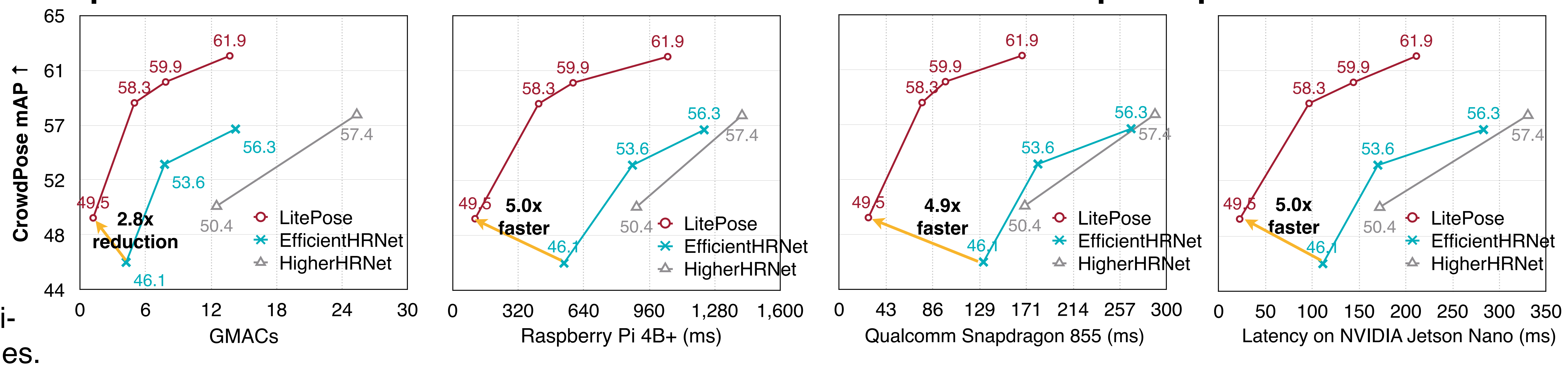
Light-weight Fusion Deconv Head





We employ the lightweight fusion deconv head to enable multi-resolution feature fusion without heavy high-resolution branches.

Compare with SOTA on the CrowdPose Dataset: 2.8-5x measured speedup


CrowdPose mAP ↑
2.8x reduction
LitePose
EfficientHRNet
HigherHRNet
GMACs
5.0x faster
Raspberry Pi 4B+ (ms)
4.9x faster
Qualcomm Snapdragon 855 (ms)
5.0x faster
Latency on NVIDIA Jetson Nano (ms)