**Simple and Lightweight Human Pose Estimation**

**Link-** [**https://arxiv.org/pdf/1911.10346.pdf**](https://arxiv.org/pdf/1911.10346.pdf)

**The 7 W's**

Q1. Most existing methods tend to pursue higher scores using complex architecture or computationally expensive models on benchmark datasets, ignoring the deployment costs in practice. In this paper, we investigate the problem of simple and lightweight human pose estimation.

Q2. It is relevant because recently, significant improvements in HPE have been achieved by using deep convolutional neural networks (DCNNs). The demand for human pose estimation networks with small model size, light computation cost, and high accuracy is increasing.

Q3.

* In the last few years, significant improvements in human pose estimation have been achieved by using DCNNs. For example, Newell et al. [6] proposed a Stacked Hourglass Network for top-down and bottom-up inference, which becomes the dominant approach on the MPII benchmark [25].
* Chen et al. [27] proposed a Cascade Pyramid Network (CPN) to refine the process of pose estimation, which is the winner of COCO 2017 keypoint challenge.
* Xiao et al. [1] provided a SimpleBaseline method that consists of a deep backbone network and several deconvolutional layers. It can achieve pretty good performance on the COCO benchmark [15]

Q4. However, these state-of-the-art methods usually involve very wide and deep networks, with numerous parameters and a huge number of floating-point operations (FLOPs). Despite top-performing, one major drawback of such complex models is that they are very time-consuming at inference time because of the heavy computation. Moreover, high memory demanding makes it extremely difficult to deploy and scale in real-world applications on edge devices.

Q5.

* We first redesign a lightweight bottleneck block with two concepts: depthwise convolution and attention mechanism. And then, based on the lightweight block, we present a Lightweight Pose Network (LPN) following the architecture design principles of SimpleBaseline [1].
* We also propose an iterative training strategy and a model-agnostic post-processing function β-Soft-Argmax to give full play to the potential of our LPN and get more accurate predicted results.
* We empirically demonstrate the effectiveness and efficiency of our methods on the benchmark dataset: the COCO key point detection dataset.
* we show the speed superiority of our lightweight network at inference time on a non-GPU platform. Specifically, our LPN-50 can achieve 68.7 in AP score on the COCO test-dev set, with only 2.7M parameters and 1.0 GFLOPs, while the inference speed is 17 FPS on an Intel i7-8700K CPU machine.

Q6. In this paper, we present a simple and lightweight network for human pose estimation, an iterative training strategy for better performance, and a post-processing function β-SoftArgmax for more accurate predictions. Our methods can achieve pretty decent results on the COCO dataset when compared with those top-performing methods, but our network is much more efficient than them in terms of inference speed.