**Ordinal Depth Supervision for 3D Human Pose Estimation**

**The 7 W's**

Q1 & Q2. The goal of this paper was to present a solution for training end-to-end ConvNets for 3D human pose estimation in the absence of accurate 3D ground truth, by using a weaker supervision signal in the form of ordinal depth relations of the joints.

Q3. Reconstruction approaches: A long line of approaches follows the reconstruction paradigm by employing 2D pose detectors to localize 2D human joints and using these lo- cations to estimate plausible 3D poses. Notably, Martinez *et al*. [25] achieve state-of-the-art results with a simple multilayer perceptron that regresses 3D joint locations, given 2D keypoints as input.

Q4. Despite the success of this paradigm, it comes with important drawbacks. No image- based evidence is used during the reconstruction step, the result is too reliant on an imperfect 2D pose detector and even for perfect 2D correspondences, the 3D estimate might fail because of the reconstruction ambiguity.

Q5.

* We propose the use of ordinal depth relations of human joints for 3D human pose estimation to bypass the need for accurate 3D ground truth.
* We showcase the flexibility of the ordinal relations by incorporating them in different network settings, where we always achieve competitive results to training with the actual 3D ground truth.
* We augment two popular 2D pose datasets (LSP and MPII) with ordinal depth annotations and demonstrate the applicability of the proposed approach to 3D pose estimation in non-studio conditions.
* We include our ordinal annotations in the training procedure of typical ConvNets for 3D human pose and exemplify their effectiveness by achieving new state- of-the-art results on the standard benchmarks.

Q6. We showcase the effectiveness and flexibility of training Convolutional Networks (ConvNets) with these ordinal relations in different settings, always achieving com- petitive performance with ConvNets trained with accurate 3D joint coordinates. Additionally, to demonstrate the potential of the approach, we augment the popular LSP and MPII datasets with ordinal depth annotations. This extension allows us to present quantitative and qualitative evaluation in non-studio conditions.